


```

Identification and cloning of new cadmium-responsive genes by
differential display
J. Biol. Chem. 273 (48), 31962-31970 (1998)
93241962
9822667
MEDLINE
PUBMED
2 (bases 1 to 907)
Liao,V.H.C., Dong,J. and Freedman,C.H.
Molecular Characterization of a Novel, Cadmium-inducible Gene from
the Nematode Caenorhabditis elegans. A NEW GENE THAT CONTRIBUTES TO
THE RESISTANCE TO CADMIUM TOXICITY
J. Biol. Chem. 277 (44), 42049-42059 (2002)
12189149
PUBMED
3 (bases 1 to 907)
Freedman,J.H. and Liao,V.H.-C.
Direct Submission
Submitted (30-AUG-2000) Nicholas School of the Environment, Duke
University, Box 90328, Durham, NC 27708-0328, USA
Derived from cosmid F35E8.
Location/Qualifiers
1..907
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/note="derived from cosmid F35E8"
1..907
/gene="cdr-1"
1..15
/gene="cdr-1"
16..849
/gene="cdr-1"
/codon_start=1
/evidence=experimental
/product="cadmium-inducible lysosomal protein CDR-1"
/protein_id="AAG30870.1"
/db_xref="GI:11120429"
/translation="MDSCLIIITAAFGAVLYLNFTVPSIKPKEDLHKDYKKOV
VTLQMRKNCNENSPGCMKIELICEFKIPYELITCTSRSENGLPVPELNGEHI
ADSLIENRSHFPIPSDELTQVAFADHLLFFVLIIRFKLWDFEYKTLIE
IIGPTFINFLMPLKAIIGKNYVNCQAGIDFELSELDEILHRLRIVENTLAKK
KFLGSEBITAADATVPSGLATVYFFRNHISDLVLEKDFPKLLEYCERVREHYVKNDF
Y"
850..907
/gene="cdr-1"
878..883
/gene="cdr-1"
/evidence=experimental
ORIGIN
Query Match 96.7%; Score 230.2; DB 3; Length 907;
Best Local Similarity 98.7%; Pred. No. 3.7e-50;
Matches 232; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 TTTTCTTTTGGAGAAATCAGCGGCGGATGCAACAGTCTTCTCAATTGGCAACT 63
Dd 673 TTTCTTTGGGAGAAATACCGGCGGATGCAACAGTCTTCTCAATTGGCAACT 732
QY 64 GTCTATATCCATTCGCAATCACATTTGGACGTTCTCGAAAGAGATTCGCCAAGTTA 123
Dd 733 GTCTATATCCATTCGCAATCACATTTGGATGTTCTCGAAAGAGATTCGCCAAGTTA 792
QY 124 TTGGATGATGTGAAGAGTTCGTATGAGTTTACCCAAAGAGCTTTACTATGTAAT 183
Dd 793 TTGGATGATGTGAAGAGTTCGTATGAGTTTACCCAAAGAGCTTTACTATGTAAT 852
QY 184 AAAATGTCAAACTAGTACATCAATAAATTTCTACGTGGCAAAAAA 238
Dd 853 AAATGTCAAACTAGTACATCAATAAATTTCTACGTGGCAAAAAA 907
RESULT 2
CEF35E8 23602 bp DNA linear INV 03-DEC-2003
LOCUS Caenorhabditis elegans cosmid F35E8, complete sequence.
DEFINITION

```

```

ACCESSION Z81529
VERSION Z81529.1
KEYWORDS GI:2653095
SOURCE HTG.
ORGANISM Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1
REFERENCE none.
AUTHORS Genome sequence of the nematode C. elegans: a platform for
TITLE Investigating biology. The C. elegans Sequencing Consortium
JOURNAL Science 282 (5396), 2012-2018 (1998)
MEDLINE 99069613
PUBMED 9851916
REMARK The C. elegans Sequencing Consortium.
REFERENCE 2 (bases 1 to 23602)
AUTHORS Gardner,A.E.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwenematode.wustl.edu
On Nov 23, 1997 this sequence version replaced gi:2546904.
Coding sequences below are predicted from computer analysis, using
predictions from GeneFinder (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
For a graphical representation of this sequence and its analysis
see: - http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
name=F35E8
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
IMPORTANT: This sequence is not the entire insert of clone F35E8.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone F35E8 is at 18616 in
sequence CEF05G11.
The true right end of clone F35E8 is at 23602 in this sequence. The
true left end of clone F05G11 is at 1324 in this sequence. The true
right end of clone F36C12 is at 104 in this sequence. The start of
this sequence (1..104) overlaps with the end of sequence CEF06C12.
The end of this sequence (23499..23602) overlaps with the start of
sequence CEF05G11.
FEATURES
Location/Qualifiers
1..23602
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/clone="F35E8"
ORIGIN
Query Match 81.7%; Score 194.4; DB 3; Length 23602;
Best Local Similarity 99.5%; Pred. No. 7.4e-41;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 30 GCGCGATGCAACAGTCTTCTCAATTGGCACTGCTATTATTCATCCGCAATCAAT 89
Dd 18625 GCGCGATGCAACAGTCTTCTCAATTGGCACTGCTATTATTCATCCGCAATCAAT 18684
QY 90 TTCGATGTTCTCGAAAGAGCTTCCCAAAGTATTGGAGTACTGTGCAAGAGTTCGTCA 149
Dd 18685 TTCGATGTTCTCGAAAGAGCTTCCCAAAGTATTGGAGTACTGTGCAAGAGTTCGTCA 18744
QY 150 TCGAGTTTACCCAAAGAGCTTTACTATGTAATTAATTTGCTCAAACTAGTACATCA 209

```

Db 18745 TGAAGTTACCAAGACCTTACTATGTGAATTAAATGTCAACTAGTACGATCA 18804

QY 210 ATAAATTTAGTGG 225

Db 18805 ATAAATTTAGTGG 18820

RESULT 3

LOCUS CEK01D12/C 36974 bp DNA linear INV 10-DEC-2003

DEFINITION Caenorhabditis elegans cosmid K01D12, complete sequence.

ACCESSION Z75543

VERSION Z75543.1 GI:1418539

KEYWORDS HTG; Potassium channel protein.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE

AUTHORS none.

TITLE Genome sequence of the nematode C. elegans: a platform for

JOURNAL investigating biology. The C. elegans Sequencing Consortium

MEDLINE Science 282 (5396), 2012-2018 (1998)

REMARK 99069613

REFERENCE The C. elegans Sequencing Consortium.

AUTHORS 2 (bases 1 to 36974)

TITLE Dobson, R.

JOURNAL Direct Submission

Submitted (29-JUN-1996) Nematode Sequencing Project, Sanger

Institute, Hinxton, Cambridge CB10 1SA, England and Department of

Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:

jes@sanger.ac.uk or rwenematode.wustl.edu

Coding sequences below are predicted from computer analysis, using

Predictions from GeneFinder (P. Green, U. Washington), and other

available information.

Current sequence finishing criteria for the C. elegans genome

sequencing consortium are that all bases are either sequenced

unambiguously on both strands, or on a single strand with both a

dye primer and dye terminator reaction, from distinct subclones.

Exceptions are indicated by an explicit note.

For a graphical representation of this sequence and its analysis

see: - [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=K01D12)

name=K01D12

IMPORTANT: This sequence is NOT necessarily the entire insert of

the specified clone. It may be shorter because we only sequence

overlapping sections once, or longer because we arrange for a small

overlap between neighbouring submissions.

This sequence is the entire insert of clone K01D12. The true right

end of clone F23H12 is at 4908 in this sequence. The start of this

sequence (1..102) overlaps with the end of sequence Z74472.

The end of this sequence (32761..36974) overlaps with the start of

sequence Z75531.

sequence Z75531.

Location/Qualifiers

1..36974

/organism="Caenorhabditis elegans"

/mol_type="genomic DNA"

/strain="Bristol N2"

/db_xref="taxon:6239"

/chromosome="v"

/clone="K01D12"

join(1282..1318,1365..1564,1640..1828,1874..2015,

2129..2505)

/genes="K01D12.1"

join(1282..1318,1365..1564,1640..1828,1874..2015,

2129..2505)

/genes="K01D12.1"

/standard_name="K01D12.1"

/note="contains similarity to Pfam domain: PF01482 (Domain

/translation="MITSMPNRRRLRAILHKKRLSKHLVAVWLISMLLINIMQDS
NEEDGVPEGTPTTFKAFYCNVFKLASRGNYEMFTSFVNJKECDGLESTKALDI
QPAVNDKVKYAFPRDENITVTGLIGHVDVAEMRLKLPKTEFFFGDMLQSPRINK
DLXEVLDGNYFQIAVSGKGMQKSYIFRKYRDATWHIGADYFFGDMQLQKPRVDIL
WMDTEGNEPVLDMHRRGPPDQGVTVICOMNIEIKOLMDKITGEREKPHDFVWKLJ
ADKKYINVRFFVYWFHQFIRTVLINVADPECTNLYN"

gene

complement (join(16035..7139,7192..7646))

/gene="K01D12.2"

complement (join(6005..7139,7192..7646))

/gene="K01D12.2"

/standard_name="K01D12.2"

/note="cDNA EST yk2e11.3 comes from this gene

cDNA EST yk2e11.5 comes from this gene

cDNA EST yk181a8.5 comes from this gene

/codon_start=1

/product="Hypothetical protein K01D12.2"

/protein_id="CAA99870.1"

/db_xref="GI:3878144"

/translation="MKSRSRQDQVDPFPPVSVLSTDSIVPECNINSTODSVPOYSY
SPATHTDILFTHSYQNNQARHGSQHHIFPHQGGOMFNFHPIHGLIACNSA
QQCSSTNSYIPIQTATVSIPTFNGEILHQSFMEFLTIDSQPIAEINIGACNSA
IQQLWTECVSIFNQLCKIQPDADNTMIANITSLPRRYIKRIFYGHRDCDYLAIEIL
EMWRSFTOTDFIVDTANDAAHOYNYVESIQPELKNSTVGNVQLDEIHESKTT
ZHPLRVQHLQNSVHTSIARTDQQLVEPTSEIHLASSTAVEIKTLNEKTAQVP
IVEIPKQNCQSGNSNTLQVPSLSVLPILPIGLAAQQLSEQLDEENPAMLET
TNEDEFELIPMETIISSENGNIDVEDSQQLSKQYQKVSLEQDQFETVETKSIIT
RINNEGSPDAMEKYGQSRIQFSSEKHSITHTYTPAESHNFGFNUHLHDHKTITQWQ
DDKHMPDGTQYQLAPRTQ"

gene

join(8020..8048,8312..8399,8466..8585,8629..8697,

8747..8888,8932..9314)

/gene="K01D12.3"

join(8020..8048,8312..8399,8466..8585,8629..8697,

8747..8888,8932..9314)

/gene="K01D12.3"

/standard_name="K01D12.3"

/note="contains similarity to Pfam domain: PF01482 (Domain

/codon_start=1

/product="Hypothetical protein K01D12.3"

/protein_id="CAA99877.1"

/db_xref="GI:3878151"

/translation="MYGKGFQYFQINVEUSPVFOAFYDCIYKPLEYKGYKEFNAF
TNWTKQDNLEVKYSMDHPASNGETKYVALPKNELLITVTLIGIHDVGAELKKE
LYKIDFADPAAPVANKDLXVENTGKGYFQAVSGESGOMQSYNDENQRYRTKE
YKADYFFNHLTKTIDMLWIDVQNEYPILQIHRENAIDKEGILLCQINVELHNGI
FGEVEDQMKLHDFVWRLIDEKRYVMKPHFVLYKMGVIRFVFMVNSDKCEADLFWN
"

gene

join(9968..10033,10136..10426)

/gene="K01D12.15"

join(9968..10033,10136..10426)

/gene="K01D12.15"

/standard_name="K01D12.15"

/note="cDNA EST yk163g9.5 comes from this gene

cDNA EST yk163g9.3 comes from this gene

cDNA EST yk206c7.3 comes from this gene

cDNA EST yk261e9.3 comes from this gene

cDNA EST yk279a12.3 comes from this gene

cDNA EST yk411a11.3 comes from this gene

cDNA EST yk206c7.5 comes from this gene

cDNA EST yk206d9.5 comes from this gene

cDNA EST yk279a12.5 comes from this gene

cDNA EST yk411a11.5 comes from this gene

cDNA EST yk552g11.3 comes from this gene

cDNA EST yk552g11.5 comes from this gene

/codon_start=1

/product="Hypothetical protein K01D12.15"

/protein_id="CAA99878.2"

/db_xref="GI:14530514"

/translation="MFSIKSVIANITFSFQVRSKRRRRTNQHLCQSGSEITKT
PILTLITLTCSEKHKQVQKQPHLRRLQCPHQLLOHLPQLESTNRSRR
PFRPMNRTSRDSN"

gene

complement (join(11177..11444,11518..11592,11643..11801,

CDS	11987..12085,12159..12481,13776..13919,14219..14302}} /gene="twk-14" complement(join(11177..11444,11518..11592,11643..11801,11987..12085,12159..12481,13776..13919,14219..14302)) /gene="twk-14" /standard name="K01D12.4" /note="similar to potassium channel protein cDNA EST YK743d10.5 comes from this gene cDNA EST YK743d10.3 comes from this gene" /codon_start=1 /product="C. elegans TWK-14 protein (corresponding sequence K01D12.4)" /protein_id="CAA99871.1" /db_xref="GI:3878145"	Db	34736 GTTTATTATCCATTTTACATCACTCTACTGCTGTTGTTGATGATTTCCAGATT 34737
gene	VLICILIVYAFGALLPHMLENEVDRIALDNRMADYQKYCKKPLNCPDEEM VRFISDGTGGLNLSRPFHLGLSFPATVISTGFTSTPHLGRITIVYGVWG CTCWLFLNLFUETVGTMSVILSRKIRYLRKSGNKPFTLLNNEDSSSS CGHMDNRVSVYKVFIFLMCLIVLTASAGIYVWENYIDSLYFISFATIG GDYVSNQDVTMRSPDLRYFVFNCLTLGACFFYCLSNVSVIVRQLLMMKMDVK VEDRSFLCFKERYMGLGRLPPKGISNLGAGCCIEHLNVFRIRYDI" join(21543..22227,22273..22403) /gene="K01D12.5" -oin(21543..22227,22273..22400) /gene="K01D12.5" /standard name="K01D12.5" /note="cDNA EST OSTF168B6_1 comes from this gene cDNA EST OSTR168B6_1 comes from this gene" /codon_start=1 /product="Hypothetical protein K01D12.5" /protein_id="CAA99872.2" /db_xref="GI:31441821"	CDS	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180 34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680
gene	SEQQTSRPNVQPOKQSVLQVLPESRNYQVQYLZTHGYNSNIQTGQP YQMGSAPGNVYSGNESYGVQAPVTAAPVHFESRQSYAAATPSPSE TERENNVRAPGHAPQAQGRSPFPFPFNPXVLSYGAAMPFLANLPQAQV YGAPIDSSHRGVDGFSLSAPAPPSSASVDJTWALRASIDANTRRHRIF" complement(join(22649..22801,22850..23341,23437..23753,23798..24207,24280..24377,24423..24458)) /gene="K01D12.6" complement(join(22649..22801,22850..23341,23437..23753,23798..24207,24280..24377,24423..24458)) /gene="K01D12.6" /standard name="K01D12.6" /note="cDNA EST cm1496 comes from this gene cDNA EST YK384f1.3 comes from this gene cDNA EST YK384f1.5 comes from this gene cDNA EST YK1046e01.5 comes from this gene cDNA EST YK1046e01.3 comes from this gene cDNA EST OSTF196H4_1 comes from this gene cDNA EST OSTR196H4_2 comes from this gene cDNA EST OSTR196H4_1 comes from this gene" /codon_start=1 /product="Hypothetical protein K01D12.6" /protein_id="CAA99873.3" /db_xref="GI:24817517"	Db	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180 34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680
CDS	ROKRAKVPQSIIEIGKADIIIVPSRQSGRMRPSPQSGERSLRNSRQ SQRNMPTEADITSGBELCTKTEKVNHNLEWNRSSKQMSRSIKTEGLKGLK OVEDIVLLVPMQEDFDNMEFVSQSVGAGSVTGRSRTLSVSDSPSFAVER FACIDIDLFVSDANLDRKELRFLDEMQALNGEVYKRSRMEPCKDSBTPAAK LYCVRAUTNALDKDEHRVWAKGRTLLADFRHTKQDPVAFNADLEWTSNDR NEEQLRQVREGRCVGTGYDAIDFIILDAFELKSPSPSVSVTKYKPNYSVMKYS /translation="MGRQSLALLSILIGETIANEYDGLSSHQSHQSVTSAPHV SEQQTTSRPNVQPOKQSVLQVLPESRNYQVQYLZTHGYNSNIQTGQP YQMGSAPGNVYSGNESYGVQAPVTAAPVHFESRQSYAAATPSPSE TERENNVRAPGHAPQAQGRSPFPFPFNPXVLSYGAAMPFLANLPQAQV YGAPIDSSHRGVDGFSLSAPAPPSSASVDJTWALRASIDANTRRHRIF" complement(join(22649..22801,22850..23341,23437..23753,23798..24207,24280..24377,24423..24458)) /gene="K01D12.6" complement(join(22649..22801,22850..23341,23437..23753,23798..24207,24280..24377,24423..24458)) /gene="K01D12.6" /standard name="K01D12.6" /note="cDNA EST cm1496 comes from this gene cDNA EST YK384f1.3 comes from this gene cDNA EST YK384f1.5 comes from this gene cDNA EST YK1046e01.5 comes from this gene cDNA EST YK1046e01.3 comes from this gene cDNA EST OSTF196H4_1 comes from this gene cDNA EST OSTR196H4_2 comes from this gene cDNA EST OSTR196H4_1 comes from this gene" /codon_start=1 /product="Hypothetical protein K01D12.6" /protein_id="CAA99873.3" /db_xref="GI:24817517"	QY	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180 34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680
Query Match	40.1%; Score 95.4; DB 3; Length 36974; Best Local Similarity 71.2%; Pred. No. 8e-15; Matches 126; Conservative 0; Mismatches 51; Indels 0; Gaps 0;	QY	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180 34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680
Db	4 TTTTITTTTGGAGGAATACCGCGGGCGGATGCAACAGCTTCTCTCAATGGCACT 63 34856 TATTATTGGAGACCAATGCTCCAGTTGATGCGAGTGTAGCGAGTGGCAGTT 34797	QY	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180 34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680
QY	64 GTCTATTATCCATTCGCGATCACAATTCGGATGTCTCCGAAAGGACTTCCCAAGTTA 123	QY	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180 34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680

Db	34736 GTTTATTATCCATTTTACATCACTCTACTGCTGTTGTTGATGATTTCCAGATT 34737
QY	124 TTGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTA 180
Db	34736 CTACGATCTGTGAAGGATCCCGAAGGAATTTATCCAAATGATTTACTATTGA 34680
RESULT 4	CEC54D10 40450 bp DNA linear INV 03-DEC-2003
LOCUS	Caenorhabditis elegans cosmid C54D10, complete sequence.
DEFINITION	275531
ACCESSION	275531.1 GI:1418465
VERSION	HTG: Serine protease inhibitor.
KEYWORDS	Caenorhabditis elegans
SOURCE	Caenorhabditis elegans
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE	1
AUTHORS	Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium Science 282 (5396), 2012-2018 (1998)
TITLE	9851916 The C. elegans Sequencing Consortium.
JOURNAL	9851916 Dobson, R.
MEDLINE	Submitted (29-JUN-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jesssanger.ac.uk or r.wane-atode.wustl.edu
REMARK	Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.
COMMENT	Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. For a graphical representation of this sequence and its analysis see: http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence? name=C54D10 IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. IMPORTANT: This sequence is not the entire insert of clone C54D10. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. The true left end of clone C54D10 is at 1 in this sequence. The true right end of clone C54D10 is at 1020 in sequence CEF28C1. The true left end of clone F28C1 is at 40353 in this sequence. The true right end of clone K01D12 is at 4214 in this sequence. The start of this sequence (1..4214) overlaps with the end of sequence CEK01D12. The end of this sequence (40353..40450) overlaps with the start of sequence CEF28C1. Location/Qualifiers 1..40450 /organism="Caenorhabditis elegans" /mol_type="genomic DNA" /strain="Bristol N2" /db_xref="taxon:6239" /clone="C54D10"
FEATURES	source
ORIGIN	Query Match 40.1%; Score 95.4; DB 3; Length 40450; Best Local Similarity 71.2%; Pred. No. 8e-15;

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPL1-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Prengen, E., Tateno, M., Cacanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genetics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RPL1-627C20; the clone sequenced to the right is AC036161. Actual start of this clone is at base position 1 of RPL1-391J13; actual end is at base position 193057 of RPL1-391J13.

FEATURES

source	Location/Qualifiers
repeat_region	1..193057
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="4"
repeat_region	/map="4"
repeat_region	/clone="RPL1-391J13"
repeat_region	/clone_lib="RPL1-11"
repeat_region	811..1283
repeat_region	/rpt_family="ERV1"
repeat_region	1373..1457
repeat_region	/rpt_family="L1"
repeat_region	2267..2555
repeat_region	/rpt_family="Alu"
repeat_region	2678..2734
repeat_region	/rpt_family="AT-rich"
repeat_region	2777..3100
repeat_region	/rpt_family="L1"
repeat_region	3098..3159
repeat_region	/rpt_family="L1"
repeat_region	3160..3193
repeat_region	/rpt_family="CA)n"
repeat_region	3194..4943
repeat_region	/rpt_family="L1"
repeat_region	8152..9376
repeat_region	/rpt_family="L1"
repeat_region	9379..10523
repeat_region	/rpt_family="L1"
repeat_region	10528..10769
repeat_region	/rpt_family="ERV1"
repeat_region	10874..10946
repeat_region	/rpt_family="MIR"
repeat_region	11034..11070
repeat_region	/rpt_family="AT-rich"
repeat_region	11242..11547
repeat_region	/rpt_family="Alu"
repeat_region	12128..12521
repeat_region	/rpt_family="ERV1"
repeat_region	13226..13946
repeat_region	/rpt_family="AT-rich"
repeat_region	14332..14441
repeat_region	/rpt_family="MaLR"
repeat_region	14753..14782
repeat_region	/rpt_family="L2"
repeat_region	15673..16020
repeat_region	/rpt_family="L1"
repeat_region	16009..16134
repeat_region	/rpt_family="L1"
repeat_region	16135..16416
repeat_region	/rpt_family="Alu"
repeat_region	16417..16963
repeat_region	/rpt_family="L1"
repeat_region	18909..18931
repeat_region	/rpt_family="CA)n"
repeat_region	18933..19089
repeat_region	/rpt_family="MaLR"
repeat_region	21867..22184
repeat_region	/rpt_family="Alu"
repeat_region	23951..24020
repeat_region	/rpt_family="TA)n"
repeat_region	24021..24244
repeat_region	/rpt_family="MER2_type"
repeat_region	24240..24685
repeat_region	/rpt_family="MER2_type"
repeat_region	26447..26542
repeat_region	/rpt_family="TATATG)n"
repeat_region	26543..26692
repeat_region	/rpt_family="L1"
repeat_region	26757..26865
repeat_region	/rpt_family="TC2"
repeat_region	27096..27177
repeat_region	/rpt_family="AT-rich"
repeat_region	27200..27801
repeat_region	/rpt_family="L1"
repeat_region	27939..28875
repeat_region	/rpt_family="L1"
repeat_region	28894..28962
repeat_region	/rpt_family="ERV1"
repeat_region	28958..29193
repeat_region	/rpt_family="ERV1"
repeat_region	29194..29604
repeat_region	/rpt_family="MaLR"
repeat_region	29605..30035
repeat_region	/rpt_family="ERV1"
repeat_region	30046..30184
repeat_region	/rpt_family="L1"
repeat_region	30195..30264
repeat_region	/rpt_family="ERV1"
repeat_region	30265..30971
repeat_region	/rpt_family="L1"
repeat_region	30970..31312
repeat_region	/rpt_family="L1"
repeat_region	31313..31606
repeat_region	/rpt_family="Alu"
repeat_region	31607..31793
repeat_region	/rpt_family="L1"
repeat_region	31820..32609
repeat_region	/rpt_family="ERV1"
repeat_region	32635..32946
repeat_region	/rpt_family="L1"
repeat_region	32959..33735
repeat_region	/rpt_family="L1"
repeat_region	33740..34204
repeat_region	/rpt_family="L1"
repeat_region	34205..34236
repeat_region	/rpt_family="CA)n"
repeat_region	34237..34403
repeat_region	/rpt_family="L1"
repeat_region	34404..34700


```

Query Match      17.24; Score 41; DB 3; Length 193057;
Best Local Similarity 52.0%; Pred. No. 1.4;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 60 AACCTGCTATTATCCATCCGAAACACATTTGGATGTTCTCGAAGAGACTTCCCAA 119
DB 52787 AACAGCAATCATATATGGGGATTCACAAATCCAAATACATTAATAGCTATCTTTAA 52728

QY 120 GTTATGAGTACTGTGAAAGAGTTCGTCAAGATTTACCCAAAGGACTTTACTATGTG 179
DB 52727 GTCAATGATTAAGACCGACATTTTTCGAAATACAGATACGATAAATTCACATAAT 52668

QY 180 AATTAATGTCACAACTAGTAGTCACATCAATATAAATCTACGTGCGAATAAATAA 236
DB 52667 TATCAAAATTAACAATCTTTTATGAAANGTACAAAAGGGAAGGTGAGAACCAAC 52611

RESULT 7
AC108193/c
LOCUS      Felis catus clone RP86-469M8, WORKING DRAFT SEQUENCE, 2 ordered
DEFINITION
ACCESSION  AC108193
VERSION     AC108193.2 GI:25815311
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Felis catus (cat)
ORGANISM    Felis catus
REFERENCE   1 (bases 1 to 138683)
AUTHORS     Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
            Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
            Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
            Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
            Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
            Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
            Margulies, E.E., Masello, C., Maskeri, B., McDowell, J.,
            Paquirian, C., Pearson, R., Portnov, M.E., Prasad, A.,
            Reddx-Dugue, R., Schandler, K., Schueler, M.G., Sison, C.,
            Stantripp, S., Thomas, J.W., Thomas, P.J., Torchman, J.W., Vogt, J.L.,
            Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
            NISC Comparative Sequencing Initiative
            Unpublished
            2 (bases 1 to 138683)
            Green, E.D.
            Direct Submission
            Submitted (26-JAN-2002) NIH Intramural Sequencing Center, 8717
            GroveMont Circle, Gaithersburg, MD 20877, USA
            3 (bases 1 to 138683)
            Green, E.D.
            Direct Submission
            Submitted (28-NOV-2002) NIH Intramural Sequencing Center, 8717
            GroveMont Circle, Gaithersburg, MD 20877, USA
            On Nov 28, 2002 this sequence version replaced gi:18376892.
            ----- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc.zoo@nhgri.nih.gov
            ----- Project Information
            Center project name: coe
            Center clone name: 469M08

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated

with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 138519 bases at least Q40
Consensus quality: 138564 bases at least Q30
Consensus quality: 138577 bases at least Q20
Insert size: 133000; agarose-fp
Insert size: 138581; sum-of-contigs
Quality coverage: 12.26x in Q20 bases; agarose-fp
Quality coverage: 11.76x in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 105452: contig of 105452 bp in length
 * 105453 105552: gap of unknown length
 * 105553 138683: contig of 33131 bp in length.

FEATURES

source

```

1. 138683
   /organism="Felis catus"
   /mol_type="genomic DNA"
   /db_xref="taxon:9685"
   /clone="RP86-469M8"
   /clone_lib="RP86"

```

misc_feature

```

1. 105452
   /note="assembly_fragment"
   /clone_end:SP6
   /vector_side:left"

```

misc_feature

```

1. 61312
   /note="clone overlaps with GenBank Accession Number
   AC108190 clone RP86-44312 (center project name cod)"

```

misc_feature

```

58366..138683
   /note="clone overlaps with GenBank Accession Number
   AC108899 clone RP86-261015 (center project name cof)"

```

misc_feature

```

105553..138683
   /note="assembly_fragment"
   /clone_end:F7
   /vector_side:right"

```

ORIGIN

```

Query Match      16.9%; Score 40.2; DB 2; Length 138683;
Best Local Similarity 52.0%; Pred. No. 2.4;
Matches 90; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 55 TTGGCACTGCTATTATCCATCCGAAACATTCGGATGTTCTCGAAGAGACTTC 114
DB 41781 TTTGCTTCTCCAGATACACTTCTCTTACAGTTAGGATAGTCAACACCTATTTC 41722

QY 115 CCAGAGTTATTTGGTACTGTCGAGAGTTGTCATGAGTTTACCCAAAGGACTTTACT 174
DB 41721 ACACACTGTGAGAGTTAAATGAGTAAATGATGATGATTAAGATTAATTAACATGTCGTCITGTT 41662

QY 175 ATGTGAATTAATTTGTCAAACTAGTAGTCAGATCAATATAAATTTCTACSTGGCA 227
DB 41661 AGTTAACTAGTAATAGTACAGTAGTAATAATAATAATAAATTTTGTCTAGGCA 41609

```

RESULT 8

AC126197/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

AC126197 250470 bp DNA linear HTG 13-MAY-2003
 Rattus norvegicus clone CH230-145B6, WORKING DRAFT SEQUENCE, 3
 unordered pieces.
 AC126197 GI:30579275
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.


```

RESULT 9
AC146595/c
LOCUS AC146595 177864 bp DNA linear HTG 06-SEP-2003
DEFINITION Mus musculus chromosome UNK clone RP24-164G18, *** SEQUENCING IN
PROGRESS ***, 6 unordered pieces.
ACCESSION AC146595
VERSION AC146595.1 GI:34494968
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 177864)
AUTHORS Wilson,R.K.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 177864)
AUTHORS Wilscot,R.K.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information
Center project name: M_B801543:8
----- Summary Statistics
Sequencing vector: M13; 0%
Chemistry: Dye-terminator ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175447 bases at least Q40
Consensus quality: 175803 bases at least Q33
Consensus quality: 175912 bases at least Q20
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1486: contig of 1486 bp in length
* 1487 1586: gap of unknown length
* 1587 2428: contig of 22695 bp in length
* 2428: gap of unknown length
* 2438: contig of 36468 bp in length
* 24382 60849: gap of unknown length
* 60850 175073: contig of 114124 bp in length
* 175074 175173: gap of unknown length
* 175174 176585: contig of 1412 bp in length
* 176586 176685: gap of unknown length
* 176686 177864: contig of 1179 bp in length.
Location/Qualifiers
1. 177864
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-164G18"
/misc_feature 1. 1486
/note="assembly_name:Contig12"
misc_feature 1587. 24281
/note="assembly_name:Contig13"
misc_feature 24382. 60849
/note="assembly_name:Contig14"
misc_feature 60950. 175073
/note="assembly_name:Contig15"
175174. 176585
/note="assembly_name:Contig7"
176686. 177864
/note="assembly_name:Contig9"

misc_feature 16.4%; Score 39; DB 2; Length 177864;
Query Match Best Local Similarity 52.8%; Pred. No. 4.9;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

ORIGIN
Query Match 16.4%; Score 39; DB 2; Length 177864;
Best Local Similarity 52.8%; Pred. No. 4.9;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 75 ATTGCGCAATCACAATTCGGATGTTCTCGAAGAGACTTCCCAAGTCTTGGAGTACTG 134
DB 145250 ATATCGAACAATCTCTAGTGGTTTGTGCAAAATTCCTCGAAGAAAGTATTAGGTTTAA 145191
QY 135 TGAAGAGATTGCTCATGAAGTTTACCCAAAGGACTTTACTATGTGAATTAATTTGCAAA 194
DB 145190 TTCACTTTTATTACTATTAATTCGCAAAATATTTAAATTTATATATTAATTAATCAAA 145131
QY 195 CTAGTAGTCAGATCAATAAATTTCTAGTGGCAAAATA 233
DB 145130 CTCATAACAATCTCTTTTGTGTTTATTATTCGCAAAATAA 145092

RESULT 10
AC108074 128635 bp DNA linear PRI 20-MAR-2002
LOCUS AC108074 Homo sapiens BAC clone RP11-717H13 from 4, complete sequence.
DEFINITION AC108074 AC024668
ACCESSION AC108074.3 GI:19551225
VERSION AC108074.3
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 128635)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 128635)
AUTHORS Trani,L., Cotton,M., Creason,K. and Nguyen,C.
TITLE The sequence of Homo sapiens BAC clone RP11-717H13
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 128635)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 128635)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 128635)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
On Mar 20, 2002 this sequence version replaced gi:18702446.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0717H13
Drafting Center: WIBR

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McSherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-384D10 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-6. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-384D10, 2000 bp overlap; the clone sequenced to the right is RP11-766D18. Actual start of this clone is at base position 21951 of RP11-384D10; actual end is at base position 128635 of RP11-717H13.

Data from AC027411 was used to finish this clone, AC108074.

FEATURES

source	Location/Qualifiers
	1..128635
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-717H13"
	/clone_lib="RPC1-11"
repeat_region	7..317
	/rpt_family="Alu"
repeat_region	400..449
	/rpt_family="L2"
repeat_region	746..811
	/rpt_family="(TA)n"
repeat_region	829..1006
	/rpt_family="(TA)n"
repeat_region	1007..1030
	/rpt_family="(CA)n"
repeat_region	1275..1708
	/rpt_family="MaLR"
repeat_region	2046..2066
	/rpt_family="(A)n"
repeat_region	3093..3133
	/rpt_family="AT_rich"
repeat_region	3735..4324
	/rpt_family="L1"
repeat_region	4833..5375
	/rpt_family="ERV1"
repeat_region	5422..6079
	/rpt_family="ERV1"
repeat_region	6486..6619
	/rpt_family="ERV1"
repeat_region	8484..8649
	/rpt_family="L2"
repeat_region	8682..9151
	/rpt_family="L1"
repeat_region	9152..10608
	/rpt_family="L1"
repeat_region	10658..10904
	/rpt_family="L2"
repeat_region	11911..12211
	/rpt_family="Alu"
repeat_region	13330..13635
	/rpt_family="Alu"
repeat_region	14025..14067
	/rpt_family="AT_rich"
repeat_region	14979..15008
	/rpt_family="(TG)n"
repeat_region	15622..15927
	/rpt_family="MaLR"
repeat_region	18551..19278
	/rpt_family="ERV1"
repeat_region	20185..20238
	/rpt_family="AT_rich"
repeat_region	20443..20932
	/rpt_family="L1"
repeat_region	20911..21112
	/rpt_family="L1"
repeat_region	21135..23419
	/rpt_family="L1"
repeat_region	24585..24745
	/rpt_family="L1"
repeat_region	24746..25029
	/rpt_family="Alu"
repeat_region	25030..25916
	/rpt_family="L1"
repeat_region	25937..26213
	/rpt_family="L1"
repeat_region	26989..27013
	/rpt_family="AT_rich"
repeat_region	27206..27471
	/rpt_family="L1"
repeat_region	27609..27719
	/rpt_family="AT_rich"
repeat_region	30364..30389
	/rpt_family="(TTG)n"
repeat_region	32127..32803
	/rpt_family="L1"
repeat_region	32843..34316
	/rpt_family="L1"
repeat_region	34453..34952
	/rpt_family="L1"
repeat_region	35071..35154
	/rpt_family="ACHobo"
repeat_region	35233..35457
	/rpt_family="L1"
repeat_region	35968..36083
	/rpt_family="L1"
repeat_region	37216..37395
	/rpt_family="MER2_type"
repeat_region	37398..37520
	/rpt_family="Alu"
misc_feature	37914..38431
	/note="match to EST AA809063 (NID:g2878469) nw17a12.s1"
repeat_region	39777..41891
	/rpt_family="L1"
repeat_region	41892..42231
	/rpt_family="L1"
repeat_region	42388..43112
	/rpt_family="L1"
repeat_region	43108..43308
	/rpt_family="L1"
repeat_region	43309..43381
	/rpt_family="MER2_type"
repeat_region	43382..45504
	/rpt_family="L1"
repeat_region	46562..46650

```

repeat_region      /rpt_family="CP-rich"
                    46753..46938
repeat_region      /rpt_family="(PATATG)n"
                    46884..47008
repeat_region      /rpt_family="(TA)n"
                    48559..48780
repeat_region      /rpt_family="Li"

```

```

Query Match      16.3%; Score 38.8; DB 9; Length 128635;
Best Local Similarity 55.1%; Pred. No. 5.7;
Matches %; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 81 CAAATCATTTCGGATGTTCTCCGAAAGGACTTCCCAAGTATTGGAGTACTGCTGAAG 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110386 CAAATCAATTTCAACATATTCGAGAGAGGATTCAGAAATTCATCAGAGAAATTAACAG 110445

QY 141 AGTTCGATGATGATTCCTCCAAAGGACTTTACTATGTAATTAATTTGCAACTATGA 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119446 AGATTAATAATATTTTAAATAACAGCAGGAAATCTGGAGCCAAATGTTCAACTGGGA 110505

QY 201 GTCAGATCAATAAAATTC 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110506 AAATGAATAAATAATAC 110523

```

```

RESULT 11
AP004700/c
LOCUS      Oryza sativa (japonica cultivar-group) DNA linear HTG 21-MAR-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone P0518H02,
            *** SEQUENCING IN PROGRESS ***.
ACCESSION  AP004700
VERSION     AP004700.1 GI:18447959
KEYWORDS   HTG; HTGS PHASE2.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Cryzeae; Oryza.

```

```

REFERENCE  1. Sasaki, T., Matsumoto, T. and Yamamoto, K.
            Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC
            clone: P0518H02
JOURNAL    Published Only in Database (2002)
REFERENCE  2. (bases 1 to 138772)
AUTHORS    Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE      Direct Submission
JOURNAL    Submitted (30-JAN-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program, Kamondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            [E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
            Tel: 81-298-38-7441, Fax: 81-298-38-7468]
COMMENT    NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

```

```

FEATURES
            Location/Qualifiers
            source
            1. .138772
               /organism="Oryza sativa (japonica cultivar-group)"
               /mol_type="genomic DNA"
               /cultivar="Nipponbare"
               /db_xref="taxon:39947"
               /chromosome="8"
               /clone="P0518H02"

```

ORIGIN

```

Query Match      15.3%; Score 38.8; DB 2; Length 138772;
Best Local Similarity 60.4%; Pred. No. 5.6;

```

```

Matches 64; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 133 TGTGAAGAGAGTTGCTGATGAAGTTTACCAAGAGACTTTACTATGTAATTAATTTGCA 192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70260 TGTCAACATTTTGAACGGGGAATATAAGGATTTTAACCTTTTACTTACTACTGTTT 70201

QY 133 AACTAGTAGTCAGATCAATTAATTTCTAGTGGCCAAAAA 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70260 AACCATTTGTTTATTTAAAAAATTATACAAATATAAAAAATAAAA 70155

```

```

RESULT 12
AP005757/c
LOCUS      Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 8,
DEFINITION PAC clone: P0413H11, complete sequence.
ACCESSION  AP005757
VERSION     AP005757.2 GI:35764583
KEYWORDS   HTG.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.

```

```

REFERENCE  1. Sasaki, T., Matsumoto, T. and Katayose, Y.
            Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC
            clone: P0413H11
JOURNAL    Published Only in Database (2002)
REFERENCE  2. (bases 1 to 147655)
AUTHORS    Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE      Direct Submission
JOURNAL    Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program, Kamondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            [E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
            Tel: 81-298-38-7441, Fax: 81-298-38-7468]
COMMENT    On Sep 25, 2003 this sequence version replaced gi:23200616.
            The orientation of the sequence is from 17 to 56 of the PAC clone.

```

FEATURES

source

```

            1. .147655
               /organism="Oryza sativa (japonica cultivar-group)"
               /mol_type="genomic DNA"
               /cultivar="Nipponbare"
               /db_xref="taxon:39947"
               /chromosome="8"
               /clone="P0413H11"

```

ORIGIN

```

Query Match      16.3%; Score 38.8; DB 8; Length 147655;
Best Local Similarity 60.4%; Pred. No. 5.6;
Matches 64; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 133 TGTGAAGAGAGTTGCTGATGAAGTTTACCAAGAGACTTTACTATGTAATTAATTTGCA 192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128424 TGTCAACATTTTGAACGGGGAATATAAGGATTTTAACCTTTTACTTACTACTGTTT 128365

QY 193 AACTAGTAGTCAGATCAATTAATTTCTAGTGGCCAAAAA 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128364 AACCATTTGTTTATTTAAAAAATTATACAAATATAAAAAATAAAA 128319

```

RESULT 13

AC132361/c

```

LOCUS      Mus musculus chromosome UNK clone RP24-149E13, WORKING DRAFT
DEFINITION AC132361
            SEQUENCE, 13 unordered pieces.
ACCESSION  AC132361
VERSION     AC132361.1 GI:22657848
KEYWORDS   HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 161010)
 McPherson, J.D. and Waterston, R.H.
 The sequence of *Mus musculus* clone
 Unpublished
 2 (bases 1 to 161010)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@watson.wustl.edu
----- Project Information -----

```

Center project name: M_BB0149E13
-----
Summary Statistics
-----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-terminator ET, 0% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 153781 bases at least Q40
Consensus quality: 154797 bases at least Q30
Consensus quality: 155407 bases at least Q20

```

```

Insert size: 179200; agarose-fp
Insert size: 159810; sum-of-ctnigs
Quality coverage: 10.01 in Q20 bases; agarose-fp
Quality coverage: 10.14 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 ctnigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the ctnigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

Accession	Strain	Host	Location/Qualifiers
1	1984:	contig of 151	
2	1985	2084: gap of unknown	
3	2085	4563: contig of 203	
4	4264	4263: gap of unknown	
5	4164	9331: contig of 503	
6	9332	9473: gap of unknown	
7	9432	16764: contig of 733	
8	9433	16764: gap of unknown	
9	16765	16864: gap of unknown	
10	16865	21118: contig of 421	
11	21119	21218: gap of unknown	
12	21219	30095: contig of 883	
13	30096	30195: gap of unknown	
14	30196	40537: contig of 103	
15	40538	40537: gap of unknown	
16	40638	53634: contig of 123	
17	53635	53734: gap of unknown	
18	53735	65578: contig of 143	
19	65579	68678: gap of unknown	
20	68679	89216: contig of 203	
21	89217	89216: gap of unknown	
22	89317	108734: contig of 193	
23	108735	108834: gap of unknown	
24	108835	159127: contig of 503	
25	159128	159127: gap of unknown	
26	159228	161010: contig of 173	
27		161010: contig of 173	
28		161010	
29		/organism="Mus musculus"	
30		/mol_type="Genomic DNA"	
31		/db_xref="taxon:10090"	
32		/chromosome="YUK"	
33		/clone="RP24-149E13"	

[illegible][illegible]

ORIGIN

Query Match	16.3%	Score 38.8;	DB 2;	Length 161010;
Best Local Similarity	51.1%	Pred. No. 5.6;		
Matches	91;	Conservative	0;	Mismatches 87; Indels 0; Gaps 0;
61	ACTGCTATATTCACATTCGCCATACATTTCCGATGTTCTCGAAAAGGACTTCCCAAAG	120		
123138	ACTATTTGTTACGAGGACGCCGCTCGAATTTATGGATGATCCAGACCGGAATTC TGARG	123079		
121	TTATTCGAGTACTGTGAAAAGGTTTCGTCATCAAGTTTTCACAAAAGGACTTTTACTATGTGA	180		
123078	GACTAGCGATGATGCTTAAGGGTTCTTAGTCTAGTCTAGTGAGACATATATATGTTTAAAG	123019		
181	ATTAATTTGCAAACTAGTCTAGTCAGATCAATATAAATTTCTAGTGGCAAAAATAAAAAA	238		
123018	TGCCAAGTGTGTCCTATCAAGGAATCAATAAAACGTGGGAGAAAAGATACAAA	122961		

RESULT 14

AF147831	AF147831	1182 bp	DNA	linear	INV 12-JUL-2000
LOCUS	Stenocéphrya sp. 28S ribosomal RNA gene, partial sequence.				
DEFINITION	AF147831				
ACCESSION	AF147831.1	GI:5059219			
VERSION					
KEYWORDS	Stenocéphrya sp.				
SOURCE	Stenocéphrya sp.				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Asilomorpha; Asiloidea; Therioidea; Stenocéphrya.				

1 (bases 1 to 1182)

Yang, L., Wiegmann, B.M., Yeates, D.K. and Irwin, M.E.
Higher-level phylogeny of the Threvidae (Diptera: based
on 28S ribosomal and elongation factor-1 alpha gene sequences
Mol. Phylogenet. Evol. 15 (3): 440-451 (2000)

20318523
MEDLINE
20318523
PUBMED
2
10860652
2 (bases 1 to 1182)
REFERENCE
Yang, L., Wiegmann, B.M., Yeates, D.K. and Irwin, M.E.
AUTHORS
TITLE
Direct Submission
SUBMITTED (03-MAY-1999) Entomology, North Carolina
JOURNAL
University, Box 7613, Raleigh, NC 27695, USA

```

FEATURES
    source
    Location/Qualifiers
        1..1182
            /organism="Stenogephyra sp."
            /mol_type="genomic DNA"

```

/db xref="taxon:95105"
 <1.->1162
 /product="28S ribosomal RNA"

rRNA

ORIGIN

Query Match 16.2%; Score 38.6; DB 3; Length 1182;
 Best Local Similarity 51.4%; Pred. No. 9.7;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 34 GATGCAAGCTCTCTCTCAATGTCGCACTGCTATTATCCATTCGCCAATCACATTCG 93
 DB 224 GATGCAAGCTCTCTCTCAATGTCGCACTGCTATTATCCATTCGCCAATCACATTCG 283
 QY 94 GATGTTCTCGAAAGCACTTCCCAAGCTATTGCGAGTACTGTGAAGAGTGTGTCAAGAA 153
 DB 284 GAAATTTATAGAAATATTTTAAATAGTATTATTAATTAAGTAAATGAAATTTGGTATGAC 343
 QY 154 GTTTCACCAAGCACTTACTATGTGATTAATTTCAATTCGCAACTAGTAGTACAG 206
 DB 344 TCTTCACTAGGTATGATCCCAATTCAGGACATTCGCCAGGTGGAGTTTGA 396

RESULT 15

AC108190/c 132389 bp DNA linear HTG 19-NOV-2002
 LOCUS Felis catus clone RP86-44312, WORKING DRAFT SEQUENCE.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 132361 bases at least Q40
 Consensus quality: 132381 bases at least Q30
 Consensus quality: 132389 bases at least Q20
 Insert size: 134000; agarose-fp
 Insert size: 132389; sum-of-contigs
 Quality coverage: 11.10x in Q20 bases; agarose-fp
 Quality coverage: 11.23x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 132389: contig of 132389 bp in length.

FEATURES

source

1..132389
 /organism="Felis catus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9685"
 /clone="RP86-44312"
 /clone_lib="RP86"

misc_feature

1..132389
 /note="assembly_fragment
 clone end:17
 vector side:left
 clone end:Sp6
 vector side:right"

misc_feature

1..70573
 /note="clone overlaps with GenBank Accession Number
 AC108196 clone RP86-591N22 (center project name ccf)"
 70580..132389

misc_feature

/note="clone overlaps with GenBank Accession Number
 AC108193 clone RP86-469M8 (center project name ccf)"
 129445..132389

misc_feature

/note="clone overlaps with GenBank Accession Number
 AC108899 clone RP86-261015 (center project name ccf)"

ORIGIN

Query Match 16.2%; Score 38.6; DB 2; Length 132389;
 Best Local Similarity 51.4%; Pred. No. 6.4;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 55 TTGGCAAGCTCTCTCTCAATGTCGCACTGCTATTATCCATTCGCCAATCACATTCG 114
 DB 112844 TTGGCTTCTCAGATACATCTTCTGTTACAGATTATGGATAGTAAACACCATTTTC 112785
 QY 115 CCAAGTATTGGAGTACTGTGAAGAGTTCGTCATGAAGTTTACCCAGAGCTTACT 174
 DB 112784 ACACACTGTGAGGTTAAATGAGGTAAATGATGATTAAGATTAATTAATTAATTTGCTTAGGCA 112725
 QY 175 AGTCAATTAAATTTGTCAACTAGTACATCAATTAATTAATTTGCTTAGGCA 227
 DB 112724 AGTCAAGTAAATGTCAGTAACTAGTACATCAATTAATTAATTTGCTTAGGCA 112672

Search completed: July 7, 2004, 15:39:07
 Job time : 1988 secs

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 13:19:03 ; Search time 2509 Seconds
(without alignments)

2832.682 Million cell updates/sec

Title: US-09-437-450A-40

Perfect score: 238

Sequence: 1 tttttttttttggaggga.....tagtggcaaaaaaaaaa 238

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_ptg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	238	100.0	238	9 AF071356	AF071356 AF071356
2	230.2	96.7	933	9 AF071357	AF071357 AF071357
3	223.2	93.8	238	9 AF071379	AF071379 AF071379
4	160.2	67.3	217	9 AF071398	AF071398 AF071398

5	159	66.8	214	9 AF071362	AF071362 AF071362
6	127	53.4	511	12 BI175173	BI175173 OSTR010H5
7	91	38.2	583	14 CB396206	CB396206 OSTR168D4
8	89.2	37.5	300	9 AUI15788	AUI15788 AUI15788
9	82	34.3	275	14 T01325	T01325 WST02646 E
10	78.4	32.9	567	14 CB396141	CB396141 CSTR167C8
11	78.4	32.9	567	14 CB396151	CB396151 OSTR167C8
12	77.2	32.4	574	14 CB397715	CB397715 OSTR193H1
13	77.2	32.4	574	14 CB397765	CB397765 CSTR193H1
14	43	18.1	574	29 CE0511398	CE0511398 tigr-gss-
15	40.8	17.1	636	29 CE717539	CE717539 tigr-gss-
16	40.8	17.1	636	29 CE717539	CE717539 tigr-gss-
17	39.6	16.6	462	28 AQ862746	AQ862746 nbeb00190
18	38.6	16.2	663	14 CB443694	CB443694 694795 MA
19	38.6	16.2	695	14 CB444053	CB444053 695179 MA
20	38	16.0	581	14 CF587970	CF587970 USDA-PP-1
21	38	16.0	1201	13 BX353711	BX353711 BX353711
22	37.8	15.9	686	28 AZ512342	AZ512342 LM0357L17
23	37.6	15.8	300	9 AA007439	AA007439 zh98b12.8
24	37.6	15.8	325	28 BZ876771	BZ876771 CH240.278
25	37.6	15.8	455	28 AZ015157	AZ015157 RPCI-23-3
26	37	15.5	358	13 BX569552	BX569552 BX569552
27	37	15.5	632	14 CB431533	CB431533 607523 MA
28	37	15.5	695	14 CB430835	CB430835 606755 MA
29	37	15.5	953	29 CC995987	CC995987 ZUBAH70TV
30	36.6	15.4	354	13 FY302654	FY302654 FY302654
31	36.2	15.2	412	12 BI863409	BI863409 r104f09.Y
32	36.2	15.2	440	13 PQ837948	PQ837948 r132h02.Y
33	36	15.1	900	29 CG767772	CG767772 Tc850.2.D
34	35.8	15.0	622	14 CA589048	CA589048 hab53b06
35	35.6	15.0	599	14 CF440904	CF440904 EST677249
36	35.4	14.9	332	14 CB393253	CB393253 OSTR116A3
37	35.4	14.9	590	28 BH186791	BH186791 C32_K_23-
38	35.4	14.9	590	29 CNS07937	CNS07937 T3 end of
39	35.4	14.9	1069	13 BQ899371	BQ899371 AGENCOURT
40	35.2	14.8	427	13 BQ602354	BQ602354 MI-P-HO-a
41	35.2	14.8	432	12 BMS37674	BMS37674 hab5c01.g
42	35.2	14.8	434	9 AA532321	AA532321 QEST.451
43	35.2	14.8	492	28 BH312146	BH312146 CH230-101
44	35.2	14.8	681	13 BX676706	BX676706 BX676706
45	35.2	14.8	726	28 BH979578	BH979578 od92908.

ALIGNMENTS

RESULT 1	AF071356	AF071356	238 bp	mRNA	linear	EST 25-NOV-1998
LOCUS	AF071356	AF071356	mRNA	from cadmium-responsive gene	Caenorhabditis elegans	
DEFINITION	AF071356	AF071356	cDNA clone DDRT16, mRNA sequence.			
ACCESSION	AF071356	AF071356.1	GI:3265101			
VERSION	AF071356	AF071356.1	GI:3265101			
KEYWORDS	AF071356	AF071356.1	GI:3265101			
SOURCE	AF071356	AF071356.1	GI:3265101			
ORGANISM	AF071356	AF071356.1	GI:3265101			
REFERENCE	AF071356	AF071356.1	GI:3265101			
AUTHORS	AF071356	AF071356.1	GI:3265101			
TITLE	AF071356	AF071356.1	GI:3265101			
JOURNAL	AF071356	AF071356.1	GI:3265101			
MEDLINE	AF071356	AF071356.1	GI:3265101			
PUBMED	AF071356	AF071356.1	GI:3265101			
COMMENT	AF071356	AF071356.1	GI:3265101			
1	(bases 1 to 238)					
Liao, V.H.-C. and Freedman, J.H.						
Cadmium-regulated genes from the nematode Caenorhabditis elegans.						
Identification and cloning of new cadmium-responsive genes by						
differential display						
J. Biol. Chem. 273 (48), 31962-31970 (1998)						
Contact: Jonathan H. Freedman						
Nicholas School of the Environment						
Duke University						
Box 90328, Durham, NC 27708-0328, USA						
Email: jon@duke.edu						
Location/Qualifiers						
I. .238						


```
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="DDRT16"
/tissue_type="whole animal"
/dev_stage="mixed population"
/clone_lib="mRNA from cadmium-responsive gene"

ORIGIN
Query Match      100.0%; Score 238; DB 9; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.5e-45;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCTTTTGGGAGGAATCAGCGCGGATGCAACAGTCTTCTCTCAATGGCACT 758
Db 1 TTTTCTTTTGGGAGGAATCAGCGCGGATGCAACAGTCTTCTCTCAATGGCACT 758

QY 61 ACTGTCATTATCCATTCGCAATCACAATTCGGATGTTCTCGAAAGGACTTCCCAAG 120
Db 61 ACTGTCATTATCCATTCGCAATCACAATTCGGATGTTCTCGAAAGGACTTCCCAAG 120

QY 121 TTATGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGGA 180
Db 121 TTATGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGGA 180

QY 181 ATTAATTTGCAAACTAGTACGATCAATAAATTTCTAGTGGCAAAAAA 238
Db 181 ATTAATTTGCAAACTAGTACGATCAATAAATTTCTAGTGGCAAAAAA 238

RESULT 2
AF071357 933 bp mRNA linear EST 25-NOV-1998
LOCUS AF071357 mRNA from cadmium-responsive gene Caenorhabditis elegans
DEFINITION cDNA clone DDRT16FC, mRNA sequence.
ACCESSION AF071357
VERSION AF071357.1 GI:3265102
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 933)
AUTHORS Liao, V.H.-C. and Freedman, J.H.
TITLE Cadmium-regulated genes from the nematode Caenorhabditis elegans.
Identification and cloning of new cadmium-responsive genes by
differential display
J. Biol. Chem. 273 (48), 31962-31970 (1998)
JOURNAL 99041962
MEDLINE 9822667
PUBMED
COMMENT Contact: Jonathan H. Freedman
Nicholas School of the Environment
Duke University
Box 90328, Durham, NC 27708-0328, USA
Email: jhf@duke.edu.

FEATURES
source
1..933
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="DDRT16FC"
/tissue_type="whole animal"
/dev_stage="mixed population"
/clone_lib="mRNA from cadmium-responsive gene"

ORIGIN
Query Match      96.7%; Score 230.2; DB 9; Length 933;
Best Local Similarity 98.7%; Pred. No. 5e-44;
Matches 232; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTTTCTTTTGGGAGGAATCAGCGCGGATGCAACAGTCTTCTCTCAATGGCACT 63
```

```
Db 699 TTTCTTTTGGGAGGAATCAGCGCGGATGCAACAGTCTTCTCTCAATGGCACT 758
QY 64 GTCTATTATCCATTCGCAATCACAATTCGGATGTTCTCGAAAGGACTTCCCAAGTTA 123
Db 759 GTCTATTATCCATTCGCAATCACAATTCGGATGTTCTCGAAAGGACTTCCCAAGTTA 818
QY 124 TTGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGGA 183
Db 819 TTGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGGA 878
QY 184 AAATTTGCAAACTAGTACGATCAATAAATTTCTAGTGGCAAAAAA 238
Db 879 AAATTTGCAAACTAGTACGATCAATAAATTTCTAGTGGCAAAAAA 933

RESULT 3
AF071379 238 bp mRNA linear EST 25-NOV-1998
LOCUS AF071379 mRNA from cadmium-responsive gene Caenorhabditis elegans
DEFINITION cDNA clone DDRT26, mRNA sequence.
ACCESSION AF071379
VERSION AF071379.1 GI:3265124
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 238)
AUTHORS Liao, V.H.-C. and Freedman, J.H.
TITLE Cadmium-regulated genes from the nematode Caenorhabditis elegans.
Identification and cloning of new cadmium-responsive genes by
differential display
J. Biol. Chem. 273 (48), 31962-31970 (1998)
JOURNAL 99041962
MEDLINE 9822667
PUBMED
COMMENT Contact: Jonathan H. Freedman
Nicholas School of the Environment
Duke University
Box 90328, Durham, NC 27708-0328, USA
Email: jhf@duke.edu.

FEATURES
source
1..238
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="DDRT26"
/tissue_type="whole animal"
/dev_stage="mixed population"
/clone_lib="mRNA from cadmium-responsive gene"

ORIGIN
Query Match      93.8%; Score 223.2; DB 9; Length 238;
Best Local Similarity 96.6%; Pred. No. 4.2e-42;
Matches 228; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TTTTCTTTTGGGAGGAATCAGCGCGGATGCAACAGTCTTCTCTCAATGGCA 60
Db 3 TTTTCTTTTGGGAGGAATCAGCGCGGATGCAACAGTCTTCTCTCAATGGCA 62
QY 61 ACTGTCATTATCCATTCGCAATCACAATTCGGATGTTCTCGAAAGGACTTCCCAAG 120
Db 63 ACTGTCATTATCCATTCGCAATCACAATTCGGATGTTCTCGAAAGGACTTCCCAAG 122
QY 121 TTATGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGGA 180
Db 123 TTATGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGGA 182
QY 181 ATTAATTTGCAAACTAGTACGATCAATAAATTTCTAGTGGCAAAAAA 236
Db 183 ATTAATTTGCAAACTAGTACGATCAATAAATTTCTAGTGGCAAAAAA 238
```

```

RESULT 4
AF071398      217 bp  mRNA  linear  EST 25-NOV-1998
LOCUS
DEFINITION AF071398 mRNA from cadmium-responsive gene Caenorhabditis elegans
cDNA clone DDX17, mRNA sequence.
ACCESSION AF071398
VERSION AF071398.1 GI:3265143
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 217)
AUTHORS Liao,V.H.-C. and Freedman,J.H.
TITLE Cadmium-regulated genes from the nematode Caenorhabditis elegans.
Identification and cloning of new cadmium-responsive genes by
differential display
J. Biol. Chem. 273 (48), 31962-31970 (1998)
JOURNAL 99041962
MEDLINE 9822667
PUBMED
COMMENT Contact: Jonathan H. Freedman
Nicholas School of the Environment
Duke University
Box 90328, Durham, NC 27708-0328, USA
Email: jonf@duke.edu.
FEATURES
source
1..217
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="caxon:6239"
/clone="DDRT7"
/tissue_type="whole animal"
/dev_stage="mixed population"
/clone_lib="mRNA from cadmium-responsive gene"
ORIGIN
Query Match 67.3%; Score 160.2; DB 9; Length 217;
Best Local Similarity 92.2%; Pred. No. 2.4e-27;
Matches 202; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
Qy 1 TTTTCTTTTGGAGGAAATCAGCGGCGGATGCAACAGTCTCTCTCAATTGGCA 60
Db 2 TTTTCTTTTGGAGGAAATCAGCGGCGGATGCAACAGTCTCTCTCAATTGGCA 61
Qy 61 ACTGCTATTATCCATTCGCAATCATTCGGATGTTCTCGAAGGACTTCCCAAG 120
Db 62 ACTGCTAT--ATCATTCGCAATCATTCGGATGTTCTCGAAGG--CATTCCAAG 118
Qy 121 TTATGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGTA 180
Db 119 TTATGGAGTACTGTGAAGAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGTA 177
Qy 181 ATTAATGTGCAACTAGTACTGAGTCAATTAATTTCT 219
Db 178 ATTAATGTGCAACTAGTACTGAGTCAATTAATTTCT 216
RESULT 5
AF071362      214 bp  mRNA  linear  EST 25-NOV-1998
LOCUS
DEFINITION AF071362 mRNA from cadmium-responsive gene Caenorhabditis elegans
cDNA clone DDX17, mRNA sequence.
ACCESSION AF071362
VERSION AF071362.1 GI:3265107
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 214)
AUTHORS Liao,V.H.-C. and Freedman,J.H.

```

```

TITLE Cadmium-regulated genes from the nematode Caenorhabditis elegans.
Identification and cloning of new cadmium-responsive genes by
differential display
J. Biol. Chem. 273 (48), 31962-31970 (1998)
JOURNAL 99041962
MEDLINE 9822667
PUBMED
COMMENT Contact: Jonathan H. Freedman
Nicholas School of the Environment
Duke University
Box 90328, Durham, NC 27708-0328, USA
Email: jonf@duke.edu.
FEATURES
source
1..214
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="DDRT2"
/tissue_type="whole animal"
/dev_stage="mixed population"
/clone_lib="mRNA from cadmium-responsive gene"
ORIGIN
Query Match 66.8%; Score 159; DB 9; Length 214;
Best Local Similarity 95.0%; Pred. No. 4.6e-27;
Matches 208; Conservative 0; Mismatches 5; Indels 6; Gaps 4;
Qy 20 AAATCAGCGGCGGATGCAACAGTCTCTCTCAATGCAACTCTCTTATTATCAATCC 79
Db 1 AAATCATGCGGCGGATGCAACAGTCT--TCTCAATTTGGCAACTCTCTATAT--CATTC 56
Qy 80 GCAATCATTTTGGATGTTCTCGAAGAGGACTTCCCAAGTATTTGAGTACTGTGAA 139
Db 57 GCAA--CACATTTGCGATGTTCTCGAAGAGGACTTCCCAAGTATTTGAGTACTGTGAA 115
Qy 140 GAGTTCGTCATGAGTTTACCCAAAGGACTTTACTATGTTGAATTAATTTGCAAACTAGT 199
Db 116 GAGTTCGTCATGAGTTT--CCCAAGGACTTTACTATGTTGAATTAATTTGCAAACTAGT 174
Qy 200 AGTCAGATCAATAAATTTCTAGTGGCAAAAAA 238
Db 175 AGTCAGATCAATAAATTTCTAGTGGCAAAAAA 213
RESULT 6
BII175173/ 511 bp  mRNA  linear  EST 09-JUL-2001
LOCUS BII175173
DEFINITION OSTR010H5.1 AD-wmcdNA Caenorhabditis elegans cDNA similar to
P35E8.11, mRNA sequence.
ACCESSION BII175173
VERSION BII175173.1 GI:14640976
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 511)
AUTHORS Reboul,J., Vaglio,P., Tzellas,N., Thierry-Mieg,N., Moore,T.,
Jackson,C., Shin-i,T., Kohara,Y., Thierry-Mieg,D., Thierry-Mieg,J.,
Lee,H., Hitti,J., Doucette-Stamm,L., Hartley,J.L., Temple,S.P.,
Brasch,M.A., Vandenhaute,J., Lamesch,P.E., Hill,D.E. and Vidal,M.
Open-reading-frame sequence tags (OSTs) support the existence of at
least 17,300 genes in C. elegans
Nat. Genet. 27 (3), 332-336 (2001)
TITLE Open-reading-frame sequence tags (OSTs) support the existence of at
least 17,300 genes in C. elegans
JOURNAL Nat. Genet. 27 (3), 332-336 (2001)
MEDLINE 21135099
PUBMED 11242119
COMMENT Contact: Reboul J, Vaglio P
Marc Vidal Laboratory
Dana Farber Cancer Institute
44 Binney Street, Boston, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 2425
Email: Jerome_Reboul@dfci.harvard.edu

```

Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project : Contact Jerome reboul@fci.harvard.edu or philippe_vaglio@fci.harvard.edu
POLYA=No.

FEATURES

Location/Qualifiers
1..511
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmcDNA"
/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"

ORIGIN

Query Match 53.4%; Score 127; DB 12; Length 511;
Best Local Similarity 90.1%; Pred. No. 9.1e-20;
Matches 136; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 4 TTTTCTTTTGGAGGAATCACGGCGGGGATGCAACAGTCTTCTCAATGGCAACT 63
Db 151 TTTCTTTTGGAGGAATCACGGCGGGGATGCAACAGTCTTCTCAATGGCAACT 92
QY 64 GTCTATATCCATCCGCAATCACATTTCGGATGTTCTCGAAGAGTCTCCCAAGTTA 123
Db 91 GTCTATATCCATCCGCAATCACATTTCGGATGTTCTCGAAGAGTCTCCCAAGTTA 32
QY 124 TTGGAGTACTGCAAGAGTTCGTCAATGAG 154
Db 31 TTGTAGTACTGGAAGAGTTCGTCAATGAG 1

RESULT 7

CB396206/c
LOCUS CB396206 583 bp mRNA linear EST 15-MAY-2003
DEFINITION OSTR168B4_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
ACCESSION CB396206
VERSION CB396206.1 GI:30737917
KEYWORDS EST.

SOURCE

Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE

1 (bases 1 to 583)
AUTHORS Rebulio, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, K., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Endress, S.A., Jenna, S., Chevet, E., Papasotiropoulos, V., Tolias, P., Ptacek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E. and Vidal, M.
C. elegans ORFome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression

TITLE

Mac. Genet. (2003) In press
Contact: Vidal M

JOURNAL

COMMENT

Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@fci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project : Contact david_hill@fci.harvard.edu or marc_vidal@fci.harvard.edu
POLYA=No.

FEATURES

Location/Qualifiers
1..583
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmcDNA"
/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"

ORIGIN

Query Match 38.2%; Score 91; DB 14; Length 583;
Best Local Similarity 70.8%; Pred. No. 2.2e-11;
Matches 121; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 4 TTTTCTTTTGGAGGAATCACGGCGGGGATGCAACAGTCTTCTCAATGGCAACT 63
Db 176 TATTATTGGAGACCAACATTCCTCCAGTTGATGCGACGGTTTATAGCCAGTGCAGTT 117
QY 64 GTCTATATCCATCCGCAATCACATTTCGGATGTTCTCGAAGAGTCTCCCAAGTTA 123
Db 116 GTTATTATCCATTTTACACTCACATCTCTACTGTTCTTGAAATGATTTCCCAAGATT 57
QY 124 TTGGAGTACTGCAAGAGTTCGTCAATGAGTTCCTCCCAAGAGTCTTACT 174
Db 56 CTACAGTACTGCAAGAGTTCGTCAATGAGTTCCTCCCAAGAGTTCCTTACT 6

RESULT 8

AU115788/c

LOCUS AU115788 300 bp mRNA linear EST 19-OCT-2000
DEFINITION AU115788 unpublished oligo-capped cDNA library Caenorhabditis elegans cDNA clone YK73695 3', mRNA sequence.

ACCESSION

AU115788

VERSION

AU115788.1 GI:10929355

KEYWORDS

EST.

SOURCE

Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE

1 (bases 1 to 300)

AUTHORS

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.

TITLE

A complementary view of the C. elegans genome

JOURNAL

COMMENT

Unpublished (2000)
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykoha@lab.nig.ac.jp.

FEATURES

source

1..300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK73695"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

ORIGIN

Query Match 37.5%; Score 89.2; DB 9; Length 300;
Best Local Similarity 70.6%; Pred. No. 8.1e-11;

Matches 132; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

QY 3 TTTTCTTTTGGAGGAATCACGCGCGGATGCAACAGTCTTCTCTCAATGGCAAC 62
 Db 190 TTTCTGTTTCGAGACAGATCACGCGCGGATGCAACAGTCTTCTCTCAATGGCAAC 131

QY 63 TGTCTATTTCATTCGCGAATCACATTCGATGTTCTCGAAAGAGGACTTCCCAAGTT 122
 Db 130 TGTATATTCTCTTCTCTCATCTGACCGGATGTTCTGCAAGAATTCGCCAAGT 71

QY 123 ATGGAGTACTGTCAAGAGCTTCGTCATGAAGTT-TACCCAAAGCACTTACATGTGAA 181
 Db 70 GCTGTAAATTTGGAAGAGTACGCAAGAAATCTATCTTACAGTTTACATTTGAA 11

QY 182 TTAAT 188
 Db 10 TTATTT 4

RESULT 9
 T01925/c
 LOCUS
 DEFINITION WEST02646 Early embryo, Strata gene (cat. #937007) Caenorhabditis elegans cDNA clone CERSX27, mRNA sequence.
 ACCESSION T01925
 VERSION T01925.1 GI:278406
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE
 AUTHORS FitzGerald, W.R., Kelley, J.M., Aubin, L., Goscoechea, M., McGonigle, M.G., Wu, A., Adams, N.E., Dubnick, M., Kerlavage, A.R., Venter, J.C. and Fields, C.A.
 TITLE Caenorhabditis elegans cDNAs
 JOURNAL Unpublished (1993)
 COMMENT Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerl@vtigr.org
 Seq primer: M13 Forward.

FEATURES
 source
 1..275
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /db_xref="taxon:6239"
 /clone="CERSX27"
 /clone_lib="Early embryo, Strata gene (cat. #937007)"

ORIGIN
 Query Match 34.5%; Score 82; DB 14; Length 275;
 Best Local Similarity 62.3%; Pred. No. 4.1e-09;
 Matches 144; Conservative 0; Mismatches 86; Indels 1; Gaps 1;

QY 4 TTTTCTTTTGGAGGAATCACGCGCGGATGCAACAGTCTTCTCTCAATGGCAAC 63
 Db 245 TTTCTTTCGAGACAGATCACGCGCGGATGCAACAGTCTTCTCTCAATGGCAAC 186

QY 64 GTCTATTATCCAT-TCGCGAATCACATTTGCGATGTTCTCGAAAGAGGACTTCCCAAGTT 122
 Db 185 GTGTATTACCCCTCTGTTGCGATACCAACGAGTGCTGGAAGAGATTTCCCAAAAT 126

QY 123 ATGGAGTACTGTCAAGAGCTTCGTCATGAAGTT-TACCCAAAGCACTTACATGTGAA 182
 Db 125 TCTTGAGTACTGGAAGAGTTCGTCGAAGTTCACCAAGAGGACTTACATGTGAA 182

QY 183 TAAATTTGCAAACTAGTAGTACATCAATATAATTTCTACGTGGCAAAAAA 233
 Db 65 TCAATTTTAGTAAATTTTCTTCTTCTTACCTTTTCTTGTAGCCAAATATAA 15

CB396141
 OSTR167C8_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
 CB396141
 CB396141.1 GI:30737852
 EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE
 AUTHORS 1 (bases 1 to 567)
 Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Endress, G.A., Jenua, S., Chevet, E., Papasotiropoulos, V., Tolias, P.P., Pracek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E. and Vidal, M.
 TITLE C. elegans ORFome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression
 JOURNAL Nat. Genet. (2003) In press
 COMMENT Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project; Contact david_hill@dfci.harvard.edu or marc_vidal@dfci.harvard.edu
 POLYA=No.

FEATURES
 source
 1..567
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /sex="Hermaphrodite and male"
 /tissue_type="whole animal"
 /dev_stage="mixed stage"
 /clone_lib="AD-wrmcDNA"
 /note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"

ORIGIN
 Query Match 32.9%; Score 78.4; DB 14; Length 567;
 Best Local Similarity 71.5%; Pred. No. 2e-28;
 Matches 103; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 3 TTTTCTTTTGGAGGAATCACGCGCGGATGCAACAGTCTTCTCTCAATGGCAAC 62
 Db 146 TTTCTGTTTCGAGACAGATCACACCGCGGATGCAACAGTCTTCTCTCAATGGCAAC 87

QY 63 TGTCTATTATTCGATTCGCGAATCACATTCGATGTTCTCGAAAGAGGACTTCCCAAGTT 122
 Db 86 TGTATATTATCTCTTCTCTCATCTGACCGGATGTTCTGCAAGAATTTCCCAAGT 27

QY 123 ATTGAGTACTGTCAAGAGGATTCG 146
 Db 26 GCTTGAATATTGGAAGAGTAGC 3

RESULT 11
 CB396151/c
 LOCUS

```

/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
Peripheral Blood"

```

ORIGIN

```

Query Match      17.1%; Score 40.8; DB 29; Length 636;
Best Local Similarity 48.3%; Pred. No. 12;
Matches 1-4; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 3 TTTT-----TTTGGGAGGAAATCAGCGCGGGGATGCGACAGTCTTCTCTCAATTGGCAAC 62
Db 236 TTTT-----TTTGGGAGGAAATCAGCGCGGGGATGCGACAGTCTTCTCTCAATTGGCAAC 62
QY 63 TGTCTATTATCCATTCGGCAATCACATTCGGATGTTCTCGAAAAGGACTTCCAAAGTT 122
Db 176 TAGATRTATTATTAGTAAGATGAGCTTTT-----TTTAAACAATATCCTTCBAAT 117
QY 123 ATTGGAGTACTGTGAAAGAGTTGGTCAATGAGTTTACCCAAAGGACTTTACTATGTGAAT 182
Db 116 TTCTAACTACTTTTGTAACTGCGATGACTTAACCTGGTGATAAAAGGAGTATTAAAAGTCT 57
QY 183 TAAATTGTCAACTAGTGTAGTCAATCAATAAATTTCTAGTGGCAGAAAABAAA 238
Db 56 ACCTTTCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1

```

Search completed: July 7, 2004, 16:21:09
Job time : 2513 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 11:39:02 ; Search time 349 Seconds

(without alignments)

2897.052 Million cell updates/sec

Title: US-09-437-450A-40

Perfect score: 238

Sequence: 1 tttttttttttggaggga.....tagctggcagaaaaa 238

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 5747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04: *

1: Geneseqn1980s: *

2: Geneseqn1990s: *

3: Geneseqn2000s: *

4: Geneseqn2001as: *

5: Geneseqn2001bs: *

6: Geneseqn2002s: *

7: Geneseqn2003as: *

8: Geneseqn2003bs: *

9: Geneseqn2003cs: *

10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34.6	14.5	199	5	ABV55254 Human pro
2	34.6	14.5	6036	6	ABN80125 Human che
3	33.6	14.1	671	3	AAZ27331 Soybean p
4	33.4	14.0	404	8	ACH29861 Human tes
5	33.4	14.0	1208	4	AAD16067 Soybean z
6	33.2	13.9	8781	6	ABL33686 Human imm
7	33	13.9	20693	4	ABL16740 Drosophila
8	32.8	13.8	571	6	ABX09781 M. incogn
9	32.8	13.8	678	6	ABQ65781 Arabidops
10	32.8	13.8	745	4	AAK87745 Human dig
11	32.8	13.8	745	4	AAI57549 Human col
12	32.8	13.8	745	6	ABS9726 cDNA enco
13	32.8	13.8	745	9	ADB92806 Human col
14	32.8	13.8	2630	4	ABL29872 Drosophila
15	32.8	13.8	3423	6	ABQ70819 Listeria
16	32.8	13.8	3581	4	ABL02096 Drosophila
17	32.8	13.8	6759	4	ABL29758 Drosophila
18	32.8	13.8	83391	6	ABQ67094 Human ang
19	32.6	13.7	2237	6	ABL58167 Human glu
20	32.6	13.7	6117	6	ABL34490 Human met
21	32.6	13.7	6561	6	ABN80020 Human che
22	32.6	13.7	37184	6	ABQ67077 Human ang
23	32.4	13.6	675	8	AAI57561 Human mag

24	32.4	13.6	1724	3	AAA76188
25	32.4	13.6	8733	4	ABL29304 Drosophila
26	32.4	13.6	16842	9	ADB54254 Pretrate
27	32.2	13.5	15373	6	ABL32467 Human imm
28	32.2	13.5	110000	6	ABA90521 04
29	32	13.4	2133	6	ABT07607 Human bre
30	32	13.4	12987	2	AAI91324 Arabidops
31	31.8	13.4	756	4	AAI96712 Human neu
32	31.8	13.4	2000	6	ABZ16757 Arabidops
33	31.8	13.4	2443	6	ABJ40935 Human nuc
34	31.8	13.4	6565	6	AAJ46465 Tumour su
35	31.8	13.4	6565	6	ABK31326 Signal tr
36	31.8	13.4	96597	8	ADA03023 Mouse Mbn
37	31.8	13.4	96597	9	ADB27261 Mouse Mbn
38	31.8	13.4	96597	9	ADC85503 Mouse Mbn
39	31.8	13.4	102457	8	ACH03408 3
40	31.6	13.3	264	3	AAA45239 Human col
41	31.6	13.3	778	4	AAH33523 Human col
42	31.6	13.3	858	4	AAI94027 Murine ch
43	31.6	13.3	4259	3	AAV4870 Human neu
44	31.6	13.3	5183	2	AAV04063 Schizosac
45	31.6	13.3	6404	6	ABN83979 Human ger

ALIGNMENTS

RESULT 1	
ABV55254	
ID	ABV55254 standard; cDNA; 199 BP.
XX	AC
XX	ABV55254;
DT	17-SEP-2002 (first entry)
XX	Human prostate expression marker cDNA 55245.
DE	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX	pharmacogenomic marker; gene; ss.
XX	Homo sapiens.
XX	W2000160860-A2.
PN	23-AUG-2001.
XX	20-FEB-2001; 2301WO-US005171.
XX	17-FEB-2000; 2000US-0183319P.
PR	16-MAR-2000; 2000US-0189862P.
PR	25-MAY-2000; 2000US-0207454P.
PR	09-JUN-2000; 2000US-0211314P.
PR	18-JUL-2000; 2000US-0219007P.
PR	13-DEC-2000; 2000US-0255281P.
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA	Schlegel R, Endege WO, Monahan JE;
PI	WPI; 2001-662795/76.
XX	Novel isolated nucleic acid molecule associated with cancerous state of
DR	prostate cells and correlating with presence of prostate cancer, useful
FT	for detecting presence of prostate cancer, stage of prostate cancer.
PT	Claim 1; Page 10673; 11750pp; English.
XX	The invention relates to an isolated nucleic acid molecule (I) comprising
XX	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC	specification or its complement. (I) is useful for: (a) assessing whether
CC	a patient is afflicted with prostate cancer; (b) monitoring the
CC	progression of prostate cancer in a patient; (c) assessing the efficacy
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 199 BP; 105 A; 34 C; 37 G; 22 T; 0 U; 1 Other;

Query Match: 14.5%; Score 34.6; DB 5; Length 199;
 Best Local Similarity 51.8%; Pred. No. 2.2;
 Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 150 TCAAGTTTACCCAAAGGCTTACTATGTGTAATTAATTGCAAACTAGTGTGATCA 209
 |||||
 Db 57 TCGAGGATACCAATTTGTCGCAAAAGGAAATATTAAGCAAACTGGTATTCAAAA 116
 |||||
 QY 210 ATAAATTTCTACGTGCGCAAAAAA 238
 |||||
 Db 117 AAAAAA 145

RESULT 2

ABN80125/c
 ID ABN80125 standard; DNA; 6036 BP.

XX AC ABN80125;

XX DT 15-JUL-2002 (first entry)

XX DE Human chemically modified disease associated gene SEQ ID NO 142.

XX KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
 KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
 KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
 KW antidiabetic; cytostatic; anticonvulsant; ds.

XX OS Homo sapiens.
 XX OS Synthetic.

XX PN WO200200927-A2.

XX PD 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007536.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI O'Leak A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-130908/17.

XX Novel nucleic acid useful for diagnosis and therapy of diseases
 PT associated with development genes such as diabetes, comprises a sequence
 PT of a segment of chemically pretreated DNA of genes associated with
 PT development.

XX C-aim 1; SEQ ID NO 142; 27pp; English.

XX The invention relates to a nucleic acid (I) comprising a sequence at
 CC least 18 bases in length of a segment of chemically pretreated DNA (II)
 CC of genes associated with development selected from 87 genes listed in the
 CC specification such as ACCPN, ADFN, or ARD1 and comprising one of 350
 CC sequences (ABN7994-ABN8033) or their complements. The invention is
 CC useful for the diagnosis or therapy of diseases associated with
 CC development genes, in particular disease related to homeobox containing
 CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
 CC associated with congenital heart disease, epilepsy, diseases related to
 CC histone deacetylation, Currarino syndrome, diseases related with the
 CC development of the brain and limb girdle muscular dystrophy and dwarfism.

CC Oligomers specific to each of the genes are useful for detecting the
 CC methylation state of all CpG dinucleotides within the 350 sequences or
 CC (II) and their complementary sequences, as primer oligonucleotides for
 CC the amplification of the 350 sequences, (III) and/or their complements and
 CC as oligomer probes for detecting the cytosine methylation state and/or
 CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
 CC patent did not form part of the printed specification but is based on
 CC sequence information supplied to Derwent by the European Patent Office
 XX
 SQ Sequence 6036 BP; 1473 A; 98 C; 1516 G; 2949 T; 0 U; 0 Other;

Query Match: 14.5%; Score 34.6; DB 6; Length 6036;
 Best Local Similarity 54.3%; Pred. No. 6.4;
 Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 110 ACTTCCCAAGTTATTCGAGTACTGTGAAGAGTTTCGTCTGAAGTTTACCCAAAGGACT 169
 |||||
 Db 3891 AATTCCTCAATTTATCTACTATTTTAACTAATTTAAATTCACCTCTCTCTTAACC 3832
 |||||

QY 170 TTACTATGTAATTAATTTGTCAACTAGTCTAGTCTAGATCAATTAATTTCTACGTGGCAA 229
 |||||
 Db 3831 TCAATTTATTCCTTAATTAATAAACAATAAATAAATAAATAAATTTCTCCCGAATA 3772
 |||||

QY 230 AAAAAAAA 238
 |||||

Db 3771 AAAAAAAA 3763

RESULT 3

AA27331
 ID AA27331 standard; cDNA; 671 BP.

XX AC AA27331;

XX DT 08-AUG-2000 (first entry)

XX DE Soybean putative glutamine amidotransferase gene 2.

XX KW Glutamine amidotransferase; histidine biosynthesi;s; herbicide; fungicide;
 KW soybean; ss.

XX OS Glycine max.

XX FH Location/Qualifiers

XX CDS 2..427

XX FT /tag= a

XX FT /product= "glutamine amidotransferase"

XX FT /partial

XX FT /transl_except= (pos:275..277,aa:Xaa)

XX FT /note= "Xaa=unknown"

XX PN WO2002026053-A2.

XX PD 18-MAY-2000.

XX PF 04-NOV-1999; 99WO-US025950.

XX PR 05-NOV-1998; 98US-0107275P.

XX PA (UJPO) DU PONT DE NEMOURS & CO E I.

XX PI Allen SM, Huang LL, Falco SC, Rafalski AJ;

XX DR WPI; 2000-376564/32.

XX DR P-PSDB; AAY94230.

XX Novel polynucleotides encoding plant glutamine amidotransferase homologs,
 PT useful for producing transgenic plants and as probes or primers.

XX PS Claim 2; Page 50; 53pp; English.

XX The present sequence is a Glycine max (soybean) putative glutamine
 CC amidotransferase coding sequence. The protein forms a crucial stage in

CC the histidine biosynthesis pathway. The sequence can be used to create
CC transgenic plants which express different amounts of the protein, to
CC identify loss of function mutants and to produce the protein in a host
CC cell, for example a bacterium. The protein can be used to identify
CC inhibitors which may be useful as fungicides and herbicides. The gene was
CC identified by computer screening for sequences which could encode
CC histidine biosynthetic enzymes

XX
SQ Sequence 671 BP; 217 A; 104 C; 144 G; 205 T; 0 U; 1 Other;

Query Match 14.1%; Score 33.6; DB 3; Length 671;
Best Local Similarity 59.4%; Pred. No. 6.4; Indels 0; Gaps 0;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 143 TTGTCATCAAGTTTACCCAAAGACCTTTACTATGTGAATTAATTTGCAAACTAGTAGT 202
DB 572 TTGTGAATATTTTAAGGAATATATTTTACCACTGTGAGCACTATCTCGACATTTCT 631

QY 203 CAGATCAATAAATTTAGTGTGCAAAAAA 238
DB 632 TATTACAAATATTTTCATGTGTCAAAAAA 667

RESULT 4

ACH29861
ID ACH29861 standard; cDNA; 404 BP.

AC ACH29861;

DT 13-OCT-2003 (first entry)

XX Human testis cDNA #247.

DE Human; ss; sequencing by hybridisation; SH; expressed sequence tag; EST;
KW Genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

PD 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRNA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 17073; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
XX 36243 cDNA sequences, appearing as ACH12769-ACH50831, whose sequence was
XX determined by the technique of SH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,

CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX
SQ Sequence 404 BP; 143 A; 43 C; 77 G; 137 T; 0 U; 4 Other;

Query Match 14.0%; Score 33.4; DB 8; Length 404;
Best Local Similarity 51.8%; Pred. No. 6.3;
Matches 100; Conservative 0; Mismatches 91; Indels 2; Gaps 1;

QY 46 TTCTCTCAATGGCACTGCTCTATTATCCATTCCTCCCAATCATCGAGTUTTCGAA 105
DB 161 TTGTCTATATGCACAGAGCTCTTTTATTAATACTAAGCAAAATTAATTTTCTGTACTAGAA 220

QY 106 AAGGACTTCCCAAGTTATTCGAGTACTGTGAAGAGTTCTCATGAAGTTTACCCAAAG 165
DB 221 AAAAATTGAACTTTTATTTAGTCTTGGCTTATAAAATGTTAATTCAGAA--TTAGTTTAT 278

QY 166 CACTTTACTATGTGAATTAATTTCTCAACTAGTAGTCAGATCAATAAAATTTACGTGG 225
DB 279 GCCTTAATTAACACTTAATTAAGCTTTGGACACTTAAAGAGCTCTAAATTTGTTGAA 338

QY 226 CAAAAA 238
DB 339 AAAAAA 351

RESULT 5

AAAD:6067/c

ID AAAD16067 standard; cDNA; 1208 BP.

XX AC AAAD16067;

XX DT 22-NOV-2001 (first entry)

XX Soybean zinc transporter (ZnT-2) cDNA from clone se6.pk0012.h2.fis.

XX Heavy metal transporter; zinc transporter; transgenic plant; soybean;

XX ZnT-2; ss.

XX OS Glycine max.

XX PH Key Location/Qualifiers

XX CDS 1..1050

XX FT /*tag= a

XX FT /product= "Soybean zinc transporter protein"

XX FT /note= "CDS does not include start codon"

XX FT /partial

XX US6278042-B1.

XX PD 21-AUG-2001.

XX PF 14-DEC-1999; 99US-00461474.

XX PR 16-DEC-1998; 98US-0112562P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Allen SM, Rafalski JA, Sakai H;

XX DR WPI; 2001-540412/60.

XX DR P-PSDB; AAE09323.

XX New arsenic transporter polypeptides and polynucleotides, useful for
XX producing transgenic plants with altered level of heavy metal tolerance.
XX Example 4; Col 41-42; 31pp; English.

CC The invention relates to heavy metal (e.g. arsenite and zinc) transporter
CC polypeptides and polynucleotides. The polypeptides of the invention are
CC useful for producing antibodies that are useful for screening cDNA
CC expression libraries to isolate full length clones of interest. The
CC antibodies are also useful for detecting the polypeptides in situ, in
CC cells or in vitro, in cell extracts. Nucleic acid molecules of the
CC invention are useful for producing transgenic plants with altered
CC tolerance to heavy metals. They are also useful as probes for genetically
CC and physically mapping the genes that are a part of, and as markers for
CC traits linked to those genes. The present sequence is soybean zinc
CC transporter (ZnT-2) protein cDNA

XX SQ Sequence 1208 BP; 347 A; 208 C; 281 G; 372 T; 0 U; 0 Other;

Query Match 14.0%; Score 33.4; DB 4; Length 1208;
Best Local Similarity 54.5%; Pred. No. 8.8;
Matches 67; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 52 CAATGGCAACGCTTATATCCATTCGCAATACATTCGGAGTTCGAAAGGAC 111
DB 864 CAACAGCCCGCTTCAAGCTTAGTAGCATCTATCTCACGAGGTGTGTTCTCCATCAGGAC 805
QY 112 TCCCAAGTATGGAGTACTGTGCAAGAGTTCGTCATGAGTTTACCCBAAGGACTTT 171
DB 804 TCCCAAAATGTTCCGAGCATGTTGATGTTGTTCCCAACAAATTAATCTGAAAGATTTAG 745
QY 172 ACT 174
DB 744 AGT 742

RESULT 6

ABL33686/c
ID ABL33686 standard; DNA; 8781 BP.

XX ABL33686;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1659.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anianaemic; cytosine; neurotropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antihumetic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olaf A. Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

XX Claim 1; SEQ ID NO 1659; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention

XX SQ Sequence 8781 BP; 2407 A; 65 C; 2093 G; 4216 T; 0 U; 0 Other;

Query Match 13.9%; Score 33.2; DB 6; Length 8781;

Best Local Similarity 55.1%; Pred. No. 19;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 121 TTATGGAGTACTGGAAGAGTTCGTATCATCAAGTTTACCCAAAGGACTTTACTATGTGA 180
DB 131 TTATAATAAATCTTTATTTACTATATAATAAATCTTTATATAATAAATCTTTATATACTA 132
QY 181 ATTAAATGTTCAACTAGTACTAGTCAGATCAATAAAATTTCACTGCGCAAAAAA 238
DB 131 ATTATAATAACAATATATATAACATATAAACAATAAATACTACTCTTCTAAAAAATAAA 74

RESULT 7

ABL16740/c

ID ABL16740 standard; DNA; 20693 BP.

XX ABL16740;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1693.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO20011042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEXE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 1693; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABBS7202). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fip.wipo.int/pub/published_pcr_sequences

XX SQ Sequence 20693 BP; 6280 A; 4195 C; 3675 G; 6543 T; 0 U; 0 Other;

Query Match 13.9%; Score 33; DB 4; Length 20693;
 Best Local Similarity 49.2%; Pred. No. 28;
 Matches 87; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
 QY 54 ATTGGCACTGCTATTATCCATTCGCAATCCGATTCGATGCTTCGAAAGGACTT 113
 Db 11638 ATTTCCAAATGATTATCTTTGATACGTATTCACCTTTGGTGTTCCTCATATGATAT 11579
 QY 114 CCCAAGTATTGCGACTACTGTGAAGAGTTCGTCAGAGTTTACCCAAAGCACTTTAC 173
 Db 11578 TCCAAATACATTCATTTGGTCCCTCAAGGTGGTTCGTAATATCCCAATATGATATAC 11519
 QY 174 TATGCAATTAATTTGTAAGTATGATAGTTCAGATCAATTAATTTACGTCGCAAA 230
 Db 11518 CATCCATTTTAAATGAATACTACTTGTCCGGAACACAAACGTTAGATGGTCAA 11462

RESULT 8
 ABX09781/c
 ID ABX09781 standard; DNA; 571 BP.
 XX AC ABX09781;
 XX DT 22-JAN-2003 (first entry)
 XX DE M. incognita RNAi molecule #38 useful for nematode control.
 XX KW RNAi molecule; double-stranded interfering RNA; nematode control;
 KW RNA mediated interference; mRNA transcripts; nematode gene; growth;
 KW development; parasitism; reproduction; RNAi vector; mRNA translation;
 KW nematode inhibitor; agricultural industry; anti-nematode; ds.
 XX OS Meloidogyne incognita.
 XX WN W0200196584-A2.
 XX PD 20-DEC-2001.
 XX PF 12-JUN-2001; 2001WO-US018911.
 XX PR 12-JUN-2000; 2000US-0210917P.
 XX PA (AKKA-) AKKADIX CORP.
 XX PI Mushegiar AR, Taylor CG, Pettelson JS, Eroshkin AM;
 XX DR WPI; 2002-39714/18.
 XX RNA mediated interference molecule useful for disrupting cellular process
 PT in a nematode, for controlling nematodes comprises genetic regulatory
 PT sequences.

Claim 1; Page 95; 103pp; English.
 The present invention relates to RNAi (double-stranded interfering RNA or
 RNA mediated interference) molecules (nematode genes), and methods of
 using these sequences in nematode control. RNAi molecules selectively
 target mRNA transcripts of essential nematode genes. The RNAi molecules
 of the invention are useful for disrupting cellular processes in a
 nematode by contacting the nematode with a composition comprising an RNAi
 molecule. The RNAi molecules are useful for killing nematodes and/or
 inhibiting their growth, development, parasitism or reproduction and also
 for the regulation of levels of specific mRNA in nematodes. Genetic
 regulatory sequences such as promoters, enhancers and terminators can be
 used in genetic constructs such as RNAi vectors which can be used for
 nematode control. The RNAi molecules are capable of targeting and
 reducing (and, in some cases, preventing) the translation of a specific
 gene product, and can be used to reduce or prevent mRNA translation in
 any tissue of the nematode because of its ability to cross tissue and
 cellular boundaries. The RNAi molecule can be contacted with a nematode
 by soaking, injection, or consumption of a food source containing an RNAi
 molecule. The RNAi molecules can also be used as an epigenetic factor to
 prevent the proliferation of subsequent generations of nematodes, to

CC produce nematode inhibitors or RNAi in the plants, and provide new
 CC biotechnological strategies for managing nematodes under sustainable
 CC agricultural conditions. ABX09677-ABX09815 represent RNAi molecules
 CC useful for the control of nematodes
 XX Sequence 571 BP; 151 A; 142 C; 72 G; 205 T; 0 J; 1 Other;
 SQ Query Match 13.8%; Score 32.8; DB 6; Length 571;
 Best Local Similarity 64.5%; Pred. No. 10;
 Matches 49; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 153 AAGGACTTTACTATCTGCAATTAATTCGCAACTAGTAGTCAGATCAATTAATTTCTACG 222
 Db 82 AAGGACTTTAAATGTTGGTAAATTTTAAATGATTTTCTGATTAATTTATTTGTA 23
 QY 223 TGGCAAAAAA 238
 Db 22 CAAAAA 7

RESULT 9
 ABQ65781/c
 ID ABQ65781 standard; DNA; 678 BP.
 XX AC ABQ65781;
 XX DT 21-AUG-2002 (first entry)
 XX DE Arabidopsis thaliana polynucleotide SEQ ID NO 358.
 XX KW Arabidopsis thaliana; thale cress; plant; transgenic; GMO; disease;
 KW stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;
 KW insecticide; antibiotic; ds.
 XX OS Arabidopsis thaliana.
 XX PN US2002059663-A1.
 XX PD 16-MAY-2002.
 XX PF 26-JAN-2001; 2001US-00770149.
 XX PR 27-JAN-2000; 2000US-0178506P.

(GOLR/) GORLACH J.
 (ANY/) AN Y.
 (HAMI/) HAMILTON C M.
 (PRIC/) PRICE J L.
 (RAIN/) RAINES T M.
 (YUY/) YU Y.
 (RAME/) RAMEAKA J G.
 (PAGE/) PAGE A.
 (MATH/) MATHW A V.
 (LEDF/) LEDFORD B L.
 (WOES/) WOESSNER J P.
 (HAAS/) HAAS W D.
 (GARC/) GARCIA C A.
 (KRIC/) KRICKER M.
 (SLAT/) SLATER T.
 (DAVI/) DAVIS K R.
 (ALLE/) ALLEN K.
 (HOFF/) HOFFMAN N.
 (HURB/) HURBAN P.
 Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,
 Ramakia JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 Hurban P;
 WPI; 2002-479224/51.
 New nucleic acid that hybridizes to Arabidopsis thaliana sequences,
 useful e.g. for preparing transgenic plants with increased resistance or

PT altered metabolism.
XX Claim 1; SEQ ID NO 358; 40pp + Sequence Listing; English.
PS
CC The invention relates to nucleic acids (I) that hybridise under stringent
CC conditions to any of 999 sequences (ABQ65424-ABQ65422) or their
CC fragments. (I) are used to express the corresponding polypeptides (II) or
CC to produce genetically modified plant cells or transgenic plants, which
CC may have improved resistance to disease or stress, or altered
CC metabolic/biosynthetic pathways (for production of commercial,
CC nutritional or medicinal products), or generally any trait of interest,
CC or can be used to screen for biologically active agents (e.g. fungicides,
CC insecticides and antibiotics). Note: The sequence data for this patent
CC did not form part of the prior specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=999909770143
XX
SQ Sequence 578 BP; 161 A; 198 G; 110 G; 209 T; 0 U; 0 Other;

Query Match 13.8%; Score 32.8; DB 6; Length 678;
Best Local Similarity 64.5%; Pred. No. 11;
Matches 49; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 163 AAGGACTTACTAGTGTCAATTAATGTCAGTCACTAGTGCAGATCAATTAATTCACG 222
Db 83 ATGTAAGTTTATGTTGGTTGAGTTGTAATACTGCTGTTGCAATTAATTAATTCGTTG 24

QY 223 TGGCAAAAAA 238
Db 23 TAGTTAAAAA 8

RESULT 10
ID AAK87745/c
ID AAK87745 standard; cDNA; 745 BP.
XX
AC AAK87745;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen coding sequence SEQ ID NO: 61.
XX
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; ss.
XX
OS Homo sapiens.
XX
XX WO200155314-A2.
XX
ED 02-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0195123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 22-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229341P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234598P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244517P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246608P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-502630/55.
 DR P-PSDB; AAM91972.

PT Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing metastases of the
 PT digestive system, particularly cancer and cancer metastases.

XX Claim 1; SEQ ID NO 61; 966pp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
 CC system antigen of the invention

SQ Sequence 745 BP; 143 A; 189 C; 167 G; 234 T; 0 U; 12 Other;

Query Match 13.8%; Score 32.8; DB 4; Length 745;

Best Local Similarity 60.5%; Pred. No. 11;

Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 153 AGTTATCCCAAGGACTTACTATGCTGAATTAATTTCAAACTAGTAGTCAGATCAATA 212

Db 108 AGATGATTTTAGGACATTTTATTTCAAATTTTACAATTTTCAATGTTNANAAAAA 49
 QY 213 AAATTCAGTGGCAAAAAA 238
 Db 48 AAAAAA 23
 RESULT 11
 AAI57549/c
 ID AA57549 standard; cDNA; 745 BP.
 XX AC AAI57549;
 XX DT 19-OCT-2001 (first entry)
 XX DE Human colorectal cancer antigen cDNA SEQ ID NO: 13.
 XX KW Human; colorectal cancer; colorectal cancer antigen; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN WO200153350-A1.
 XX PD 02-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US001350.
 XX PR 21-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0218293P.
 PR 26-JUL-2000; 2000US-0223963P.
 PR 26-JUL-2000; 2000US-0223964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0231444P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 21-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 25-SEP-2000; 2000US-0234999P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236322P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241783P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241828P.
 PR 01-NOV-2000; 2000US-0244611P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 XX P-PSDB; AM38571.
 XX WPI: 2001-457727/49.
 DR P-PSDB; AM38571.
 XX
 PT Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the colon and rectum including colorectal cancers
 PT and also for testing and detection e.g. diagnosis.
 XX
 PS Claim 1; SEQ ID NO 13; 522pp + Sequence Listing; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of colorectal cancer antigens. These are shown in AA15747-
 CC AA157619 and AA157619-AM38641. These can be used in the diagnosis,
 CC prevention and treatment of cancer of the colon and/or rectum. The
 CC present sequence is a colorectal cancer antigen coding sequence of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 745 BP; 143 A; 189 C; 167 G; 234 T; 0 U; 12 Other;
 Query Match 13.8%; Score 32.8; DB 4; Length 745;
 Best Local Similarity 60.5%; Pred. No. 11;
 Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 QY 153 AGTTTACCCAAAGGACTTTTACTATGTGAATTAAATTCAAACTAGTAGTCAGATCAATA 212
 Db 108 AGATGATTTTAGGCGCATTTTATTTTAATTAATTTACAACTTAATGTTNANAAAAAA 49
 QY 213 AAATTCACGTGGCAAAAAA 238
 Db 48 AAAAAAAAAAAAAAAAAAAAAA 23
 RESULT 12
 ABS99726/c
 ID ABS99726 standard; cDNA; 745 BP.
 AC ABS99726;
 XX
 DT 18-DEC-2002 (first entry)
 XX
 DE cDNA encoding human colorectal cancer related protein #3.
 KW Human; colorectal cancer related protein; colon; rectum;
 KW colorectal cancer metastasis; gastrointestinal disorder; cytostatic;
 KW Gene; ss.
 XX
 CS Homo sapiens.
 XX
 PN US2002119919-A1.

XX PD 29-AUG-2002.
 XX PF 17-JAN-2001; 2000US-00764855.
 XX PR 31-JAN-2000; 2000US-0179065P.
 XX (ROSE/) ROSEN C A.
 XX (RUBEN/) RUBEN S M.
 XX (BARA/) BARASH S C.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI: 2002-731367/79.
 XX P-PSDB; ABG97623.
 XX New colorectal cancer polypeptide for diagnosing, prognosing, preventing,
 PT and treating immune, hyperproliferative, liver, kidney, reproductive
 PT disorders and for identifying modulators of therapeutic use.
 XX C:aim 4; SEQ ID NO 13; 153pp; English.
 XX The present invention relates to the isolation of novel human colorectal
 CC cancer related proteins, and polynucleotide sequences encoding them. The
 CC sequences of the invention are useful in the diagnosis, treatment,
 CC prevention and/or prognosis of the colon and/or rectum, including
 CC colorectal cancer, colorectal cancer metastases, and gastrointestinal
 CC disorders such as dysphagia, peptic esophagitis, gastric reflux,
 CC irritable bowel syndrome, and peritoneal diseases. The invention also
 CC describes antibodies that bind colorectal cancer related proteins,
 CC vectors, host cells, and recombinant and synthetic methods for producing
 CC human colorectal cancer related polynucleotides, polypeptides, and/or
 CC antibodies. ABG97624-ABG97796 represent cDNA sequences encoding human
 CC colorectal cancer related proteins. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipspidentry.html
 XX Seqdata.uspto.gov/psipspidentry.html
 XX Sequence 745 BP; 143 A; 189 C; 167 G; 234 T; 0 U; 12 Other;
 Query Match 13.8%; Score 32.8; DS 6; Length 745;
 Best Local Similarity 60.5%; Pred. NO. 11;
 Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 Qy 153 AGTTTACCCAGGACTTTACTATGTGAAATTAATTTGCAACTAGTACGATCAATA 212
 Db 108 AGATGATTTAGGACATTTATTTTAAATTTAAATTTACATCTATGCTGTAATAAAA 49
 Qy 213 AAATTTACGTCGCAAAAAA 238
 Db 48 AAAAAA 23
 RESULT 13
 ADB92806/c
 ID ADB92806 standard; cDNA; 745 BP.
 XX AC ADB92806;
 XX 04-DEC-2003 (first entry)
 XX Human colorectal cancer related polypeptide cDNA #3.
 XX ss; gene; human; colorectal cancer; antigen; gene therapy;
 KW gastrointestinal disorder; inflammatory disease; infection; cancer;
 KW small intestine neoplasm; small intestine carcinoma tumour;
 KW small intestine non-Hodgkin's lymphoma; small bowel lymphoma; ulcer;
 KW peptic ulcer; Bruton's disease; X linked infantile agammaglobulinaemia;
 KW severe combined immunodeficiency; DiGeorge anomaly;
 KW hyperproliferative disorder; acute lymphoblastic leukaemia;
 KW acute lymphocytic leukaemia; urinary system disorder; cortical necrosis;
 KW kidney infarction; cardiovascular disorder; carcinoma heart disease;
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; sinusitis;

KW musculoskeletal system disorder; Albers-Schonberg disease;
 KW Marfan's syndrome; neurological disease; phenylketonuria;
 KW Wernicke's encephalopathy; Alzheimer's disease; endocrine disorder;
 KW Grave's disease; Cushing's syndrome; reproductive system disorder;
 KW prostatitis; benign prostatic hyperplasia; benign prostatic hyperplasia;
 KW thrombosis; atherosclerosis; myocardial infarction; ischaemic attack.
 XX Homo sapiens.
 OS
 XX US2003054420-A1.
 XX 20-MAR-2003.
 XX 11-FEB-2002; 2002US-00072349.
 XX 31-JAN-2000; 2000US-0179065P.
 XX 04-FEB-2000; 2000US-0180628P.
 XX 24-FEB-2000; 2000US-0184664P.
 XX 02-MAR-2000; 2000US-0186350P.
 XX 16-MAR-2000; 2000US-0189874P.
 XX 17-MAR-2000; 2000US-0190076P.
 XX 18-APR-2000; 2000US-0198123P.
 XX 19-MAY-2000; 2000US-0205151P.
 XX 07-JUN-2000; 2000US-0209467P.
 XX 28-JUN-2000; 2000US-0211486P.
 XX 30-JUN-2000; 2000US-0215135P.
 XX 07-JUL-2000; 2000US-0216647P.
 XX 07-JUL-2000; 2000US-0216880P.
 XX 11-JUL-2000; 2000US-0217487P.
 XX 11-JUL-2000; 2000US-0217496P.
 XX 11-JUL-2000; 2000US-0218290P.
 XX 26-JUL-2000; 2000US-0220963P.
 XX 26-JUL-2000; 2000US-0220964P.
 XX 14-AUG-2000; 2000US-0224518P.
 XX 14-AUG-2000; 2000US-0224519P.
 XX 14-AUG-2000; 2000US-0225213P.
 XX 14-AUG-2000; 2000US-0225214P.
 XX 14-AUG-2000; 2000US-0225266P.
 XX 14-AUG-2000; 2000US-0225267P.
 XX 14-AUG-2000; 2000US-0225268P.
 XX 14-AUG-2000; 2000US-0225447P.
 XX 14-AUG-2000; 2000US-0225757P.
 XX 14-AUG-2000; 2000US-0225758P.
 XX 18-AUG-2000; 2000US-0225759P.
 XX 22-AUG-2000; 2000US-0226681P.
 XX 22-AUG-2000; 2000US-0226686P.
 XX 22-AUG-2000; 2000US-0227182P.
 XX 23-AUG-2000; 2000US-0227009P.
 XX 30-AUG-2000; 2000US-0228924P.
 XX 01-SEP-2000; 2000US-0229287P.
 XX 01-SEP-2000; 2000US-0229343P.
 XX 01-SEP-2000; 2000US-0229344P.
 XX 01-SEP-2000; 2000US-0229345P.
 XX 05-SEP-2000; 2000US-0229509P.
 XX 05-SEP-2000; 2000US-0229513P.
 XX 06-SEP-2000; 2000US-0230437P.
 XX 06-SEP-2000; 2000US-0230438P.
 XX 08-SEP-2000; 2000US-0231242P.
 XX 08-SEP-2000; 2000US-0231243P.
 XX 08-SEP-2000; 2000US-0231244P.
 XX 08-SEP-2000; 2000US-0231413P.
 XX 08-SEP-2000; 2000US-0231414P.
 XX 08-SEP-2000; 2000US-0232080P.
 XX 12-SEP-2000; 2000US-0231968P.
 XX 14-SEP-2000; 2000US-0232397P.
 XX 14-SEP-2000; 2000US-0232398P.
 XX 14-SEP-2000; 2000US-0232399P.
 XX 14-SEP-2000; 2000US-0232400P.
 XX 14-SEP-2000; 2000US-0232401P.
 XX 14-SEP-2000; 2000US-0233063P.

PR	08-DEC-2020; 2000US-025198SP.
PR	08-DEC-2020; 2000US-025199OP.
PR	11-DEC-2020; 2000US-025409FP.
PR	05-JAN-2001; 2001US-025967BP.
PR	17-JAN-2001; 2001US-007648SP.
XX	(HUMA-) HUMAN GENOME SCI INC.
XX	Rosen CA, Ruben SM, Barash SC;
PI	WPI; 2003-708345/67.
DR	P-PSDE; ADB92879.
DR	Novel colorectal cancer antigen useful for treating, preventing,
PT	diagnosing and/or prognosing gastrointestinal disorders, infections,
PT	cancers such as intestinal neoplasms, ulcers.
XX	Claim 3; SEQ ID NO i3; l79pp; English.
PS	The invention relates to a colorectal cancer antigen. The antigen is
CC	useful for chromosome identification, chromosome mapping, radiation
CC	hybrid mapping or gene therapy, or as hybridisation probes for
CC	differential identification of the tissues or cell types present in a
CC	biological sample. The antigen is useful for treating, preventing,
CC	diagnosing and/or prognosing gastrointestinal disorders, including
CC	inflammatory diseases and/or conditions, infections, cancers (e.g.
CC	intestinal neoplasms (carcinoma) of the small intestine, nor-
CC	Hodgkin's lymphoma of the small intestine, small bowel lymphoma)) and
CC	ulcers (e.g. peptic ulcers). The antigen and its nucleic acid are useful
CC	to provide immunological probes for differential identification of the
CC	tissue. The antigen and its nucleic acid are useful for treating,
CC	preventing, diagnosing and/or prognosing diseases, disorders and/or
CC	conditions of the immune system e.g. Bruton's disease, X linked infantile
CC	agammaglobulinaemia, severe combined immunodeficiencies, DiGeorge
CC	anomaly, etc. The antigen and its nucleic acid is useful for treating,
CC	preventing and/or diagnosing hyperproliferative disorders (e.g. acute
Query Match	13.8%; Score 32.8; DB 9; Length 745;
Best Local Similarity	60.5%; Pred. No. 11;
Matches	52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY	153 AGTTTACCACAAAGCATTCTACTGTGCAATTAATTCACAACTAGTAGTCAGATCAATA 212
Db	108 AGATGATCTTAGGGACATTTATTTAATTAATTTACAATCTCATGTGNAAAAAAA 49
QY	213 AAATTCACGTGGCAAAAAAAAAA 238
Db	48 AAAAAAAAAAAAAAAAAAAAAAAAAA 23
RESULT 14	
ABL29872	ABL29872
ID	ABL29872 standard; DNA; 2630 BP.
XC	ABL29872;
AC	
XX	26-MAR-2002 (first entry)
DT	
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 41089.
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.; gene; ds.
OS	Drosophila melanogaster.
PN	WO200171042-A2.
PP	27-SEP-2001.
PF	23-MAR-2001; 2001WO-US009233..
XX	23-MAR-2000; 2000US-C191637P.
PR	11-JULI-2000; 2000US-C0614150.

```
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers BW;
XX DR WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 41089; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2630 BP; 725 A; 579 C; 613 G; 713 T; 0 U; 0 Other;
SQ
Query Match 13.8%; Score 32.8; DB 4; Length 2630;
Best Local Similarity 49.4%; Pred. No. 17;
Matches 55; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 23 TCACGGGGGGGATGCAACAGCTCTCTCAATTGGCAACTGCTTATTATCCATTCCGCA 82
Db 505 TAATGCCAGTTAAAGCTTAAGATGTTTGTCAAGTGGTGTCAATAATTATATGTTATCGT 564
QY 83 ATCACATTTCGGATGTTCTCGAAAGAGCTTCCCAAGTATTGGAGTACTGTGAAGAG 142
Db 565 CTATCAGTATGATGATGCTTAAAGATGGTATGTTGTCATTTCAATGAGTTAAGTTAAAAA 624
QY 143 TTCGTCATGGAAGTTTACCCAAAGGACTTCTATGTAATTAAATGTCAAA 194
Db 625 TTGTGAGCTTTTGGCCCTTCGTCATGTGAAGATGATATTTCACGGGCAAA 676
RESULT 15
ABQ708.9
ID ABQ70819 standard; DNA; 3423 BP.
XX AC ABQ70819;
XX
XX 29-AUG-2003 (revised)
DT 29-AUG-2002 (first entry)
XX
DE Listeria monocytogenes 4b contig DNA sequence #761.
XX
XX Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX Listeria monocytogenes ATCC 19115.
XX
XX WO200228891-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-FR003061.
XX
XX 04-OCT-2000; 2000FR-0002697.
XX
XX (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
XX
XX WPI; 2002-332479/37.
```

```
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX
XX Claim 14; SEQ ID NO 3632; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences (ABQ67188-
CC ABQ71212) from Listeria sp. The sequences are useful as probes and
CC primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of gene
CC expression. Proteins encoded by the nucleic acid sequences can be used to
CC screen for compounds that modulate gene expression, replication and
CC pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in anti-
CC Listeria vaccines. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
CC on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 3423 BP; 1040 A; 693 C; 620 G; 1070 T; 0 U; 0 Other;
SQ
Query Match 13.8%; Score 32.8; DB 6; Length 3423;
Best Local Similarity 59.8%; Pred. No. 18;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 58 GCAACTGTCTATTATCCATTCCGCAATCCACATTTCGGATGTTCTCGAAGAGCACTCCCA 117
Db 2377 GAATTGTTTATCTTTTCCCAAGCAAAACGAATGATGATGATGATGATGATGATGATGAT 2436
QY 118 AAGTTATGAGTACTGTGGAAGAGTTCGTCA 149
Db 2437 CAACATCTGGCATAGTACGAATGTTTCGCA 2468
Search completed: July 7, 2004, 15:07:29
Job time : 353 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 13:49:08 ; Search time 81 Seconds
(without alignments)
1630.596 Million cell updates/sec

Title: US-09-437-450A-40
Perfect score: 238
Sequence: 1 ttttttttttttgggagga.....tagtggcaaaaaa 238

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Parents NA.*
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS-COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	33.4	14.0	1208	3	US-09-461-474-11
C 2	32.4	13.6	1724	3	US-09-197-679A-1
C 3	31.8	13.4	832	4	US-09-621-976-2813
C 4	31.6	13.3	5183	2	US-08-870-518-7
C 5	31.4	13.2	6124	4	US-08-213-4198-3
C 6	31.4	13.2	29793	4	US-09-302-812-38
C 7	31.4	13.2	29793	4	US-09-511-477-38
C 8	31.4	13.2	29793	4	US-09-511-507-38
C 9	31.2	13.1	696	4	US-09-227-357-69
C 10	30.8	12.9	2244	4	US-09-601-198-149
C 11	30.6	12.9	1664976	4	US-08-916-421B-1
C 12	30.4	12.8	401	4	US-09-643-597-266
C 13	30.4	12.8	401	4	US-09-480-884A-266
C 14	30.4	12.8	401	4	US-09-542-615A-266
C 15	30.4	12.8	401	4	US-09-606-421B-266
C 16	30.4	12.8	640681	4	US-09-790-988-1
C 17	30.2	12.7	738	4	US-09-552-225A-10
C 18	29.8	12.5	1269	4	US-09-322-409-99
C 19	29.8	12.5	1269	4	US-09-322-409-101
C 20	29.8	12.5	1269	4	US-09-451-527-99
C 21	29.8	12.5	1369	4	US-09-451-527-161
C 22	29.8	12.5	1302	4	US-09-322-409-91
C 23	29.8	12.5	1302	4	US-09-322-409-93
C 24	29.8	12.5	1302	4	US-09-451-527-91
C 25	29.8	12.5	1302	4	US-09-451-527-93
C 26	29.8	12.5	3981	2	US-08-955-138-2
C 27	29.6	12.4	477	1	US-08-313-608B-2

C 28	29.6	12.4	477	2	US-08-459-324-2	Sequence 2, Appli
C 29	29.6	12.4	477	4	US-09-107-532A-3599	Sequence 3599, Ap
C 30	29.6	12.4	479	1	US-08-313-608B-1	Sequence 1, Appli
C 31	29.6	12.4	479	2	US-08-459-324-1	Sequence 1, Appli
C 32	29.6	12.4	246240	2	US-08-724-394A-20	Sequence 20, Appl
C 33	29.6	12.4	246240	2	US-08-724-394A-21	Sequence 21, Appl
C 34	29.6	12.4	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 35	29.4	12.4	1465	4	US-09-444-336-6	Sequence 6, Appli
C 36	29.2	12.3	2322	3	US-09-152-406-1	Sequence 1, Appli
C 37	29.2	12.3	2322	3	US-09-152-406-5	Sequence 5, Appli
C 38	29.2	12.3	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 39	29.2	12.3	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 40	29	12.2	243	4	US-09-540-236-1400	Sequence 1400, Ap
C 41	29	12.2	999	4	US-09-134-001C-362	Sequence 362, App
C 42	29	12.2	2163	4	US-09-134-000C-2226	Sequence 2226, Ap
C 43	28.8	12.1	3338	4	US-09-489-847-117	Sequence 117, App
C 44	28.6	12.0	1287	4	US-09-833-381-1759	Sequence 1759, Ap
C 45	28.6	12.0	1464	1	US-08-292-688A-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-461-474-11/c
; Sequence 11, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant Metal Transporters
; FILE REFERENCE: BBI303 US NA
; CURRENT APPLICATION NUMBER: US/09/461,474
; CURRENT FILING DATE: 1998-12-14
; EARLIER APPLICATION NUMBER: 60/112,562
; EARLIER FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1208
; TYPE: DNA
; ORGANISM: Glycine max
US-09-461-474-11

Query Match 14.0% Score 33.4; DB 3; Length 1208;

Best Local Similarity 54.5%; Pred. No. 0.56; Mismatches 56; Indels 0; Gaps 0;
Matches 67; Conservative 0

QY	52	CAATTGGCAACTGTCTATTATCCATTCGGCAATCACATTTCGGATGTTTCGAAAGGAC	111
Db	864	CAACAGCCCTTTCAAGCTTAGCATCTATCTCAGAGGTGTGTTCTCCATCAGGAC	805
QY	112	TTCCCAAGTATTGGAGTACTGTGAAAGAGTTCGATGATGATTCACCAAGACTTT	171
Db	804	TTCCAAATGTTTCGACATGTTGATGTTGTCCTCCATAACAATTACTGAAAGATTAG	745
QY	172	ACT 174	
Db	744	AGT 742	

RESULT 2

US-09-197-679A-1
; Sequence 1, Application US/09197679A
; Patent No. 6287590
; GENERAL INFORMATION:
; APPLICANT: Foley, Patricia L
; TITLE OF INVENTION: Vaccine Against Swine Influenza
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Associates
; STREET: P.O. Box 2509

```

Dd      235 TWCYCYWYKYNRWKSKTOWSGRGGWNYSYSTSYSNWASWYOTCTCKTWGRWSTAY 176
Qy      67 TAT--TATCCATTCCGCAATCACATTTTCGGATGTTCTCGAAAAGGACITCCCCAAAGTTAT 124
Dd      175 MAWGKKWRVATTTWRRAWWWWWAAWTMMWYWWAWKCSSRGAAMYRRTMWNGYRYWNRK 116
Qy      125 TGGACTACTGTGAAGAGTGTCGTCACTGAAGTYTACCNAAGGACITTTACTATCTGAATA 184
Dd      115 KSYRETRCWAYAKWKTSYTNCRWKRWKCRCOMMMMAAAYKTMMORACWKTRTWRNA 56
Qy      185 AATTCCTCAAATAGTAGTCACATCAATAAAAAATTCCTACGTCGCAAAAAA 237
Dd      55 WAWRWMTMWWXYYYPRAMKRWKRWKRSWSNMWMAWXTRWARMWRW 3

RESULT 4
US-08-870-518-7
; Sequence 7, Application US/08870518
; Patent No. 592556
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger J.
; APPLICANT: Galcheva-Gargova, Zoya
; TITLE OF INVENTION: NON-ACTIVATED RECEPTOR COMPLEX
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,518
; FILING DATE: 06-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,219
; FILING DATE: 06-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04020/102001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-870-518-7

Query Match          13.3%; Score 31.6; DB 2; Length 5183;
Best Local Similarity 48.9%; Pred.No.3.5;
Matches            85; Conservative           0; Mismatches    89; Indels       0; Gaps        0;

Qy      50 CTCATFTGGCAACTCTATTAATCCCATTCGCATACATTCGATTCGGATGTCCTCGAAGG 109
Dd      3780 CTCCTGTGCTGTGTTCATTCATCAAAACTTCTCGTATTTCTTAAGACCTCATCAATAG 3839
Qy      110 ACTTCCCAGAATTATGGAGTACTGTGAAGAGTTTCGTCAATGAAGTTTACCCAAAGGACT 169
Dd      3840 AAGGTGTCAATCTTGTCACAGTGTGTACTCGAAGATTGACCGTTGTGCAATGGAT 3899
Qy      170 TTACTATCTGAATTAATTTGTCACAACTAGTAGTCAGATCAATAAATTTCTAAGT 223

```

Db 390C GTATTATGCGACATGAGATGGAGCTTGTGTAAGACACACTCAATGATTCCTTGT 3953

RESULT 5

US-08-213-419B-3/c

Sequence 3, Application US/08213419B

Patent No. 6333406

GENERAL INFORMATION:

APPLICANT: Inselburg, J. et al.

TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM

TITLE OF INVENTION: AND USES THEREFOR

FILE REFERENCE: J11-002CNP

CURRENT APPLICATION NUMBER: US/08/213,419B

PRIOR FILING DATE: 1994-03-14

PRIOR APPLICATION NUMBER: US 07/870,506

PRIOR FILING DATE: 1992-04-17

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 3

LENGTH: 6:24

TYPE: DNA

ORGANISM: Plasmodium falciparum

FEATURES:

NAME/KEY: CDS

LOCATION: (2407)..(2439)

NAME/KEY: CDS

LOCATION: (2598)..(3404)

NAME/KEY: CDS

LOCATION: (3580)..(3720)

NAME/KEY: CDS

LOCATION: (3850)..(5835)

US-08-213-419B-3

Query Match 13.2%; Score 31.4; DB 4; Length 6124;

Best Local Similarity 47.3%; Pred. No. 4.3; Indels 0; Gaps 0;

Matches 95; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 38 CAACAGTCTCTCTCAATGGCAACTGTCTATTATCCATTCCGCAATCAATTCGGTGG 97

DB 865 CATGAATACCTTCATTCGTAAATGGAAATTTTTCATATGATATATATATA 806

QY 98 TTCGGAAGAGACTCCCAAGTTATCTGGAGTACTGTGAAGAGTTCGTCAAGTTT 157

DB 805 TATATATTAAAGCAAAATAGATATTAAATAAAAGAAATACATAATAGTTTAA 746

QY 158 ACCCAAGAGACTTTACTATGTGAATTAATTCCTCAACTAGTAGTCAGATCAATAAAT 217

DB 745 CTAAAGACTAGTTTAAATTCATATTTTAAATTTTAAATGAAATGAAATGAAATGAA 686

QY 218 CTACGTGGCAAAAAA 238

DB 685 TAAATAAATAAATAAATAA 665

RESULT 6

US-09-302-812-38/c

Sequence 38, Application US/09302812B

Patent No. 6333148

GENERAL INFORMATION:

APPLICANT: JACOBSON, Myron K.

APPLICANT: JACOBSON, Elaine L.

APPLICANT: AM, Jean-Christophe

APPLICANT: LIN, Winston

TITLE OF INVENTION: GENES ENCODING SEVERAL POLY (ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZ

TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE

TITLE OF INVENTION: THEREWITH

FILE REFERENCE: NIAD 201

CURRENT APPLICATION NUMBER: US/09/302,812B

CURRENT FILING DATE: 1999-04-30

EARLIER APPLICATION NUMBER: 60/083,768

EARLIER FILING DATE: 1998-05-01

NUMBER OF SEQ ID NOS: 38

SEQ ID NO 38

Query Match 13.2%; Score 31.4; DB 4; Length 29793;

Best Local Similarity 51.0%; Pred. No. 7.8; Indels 0; Gaps 0;

Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 94 GATGTTCTCGAAAAGGAGCTTCCCAAGTTATTGGAGTACTGTGAAGAGTTCGTCATGAA 153

DB 21188 GATTTCTACATCTCTTATGCGCAAGCATGTTTTCATATGATGAAATTTCTGTT 21129

QY 154 GTTTACCCCAAGGACITTTACTATGTGAATTAATTTCAAACTAGTAGTCAGATCAATAA 213

DB 21128 TTTTCCCATTAAGTGAATTCATAATTTTATGTCATATGCTTGGTCTAGTAAAT 21069

QY 214 AATCTAGTCGCAAAAAA 238

DB 21068 TTTAATTTTGAATAAATAA 21044

RESULT 7

US-09-511-477-38/c

Sequence 38, Application US/09511477

Patent No. 6337282

GENERAL INFORMATION:

APPLICANT: JACOBSON, Myron K.

APPLICANT: JACOBSON, Elaine L.

APPLICANT: AME, Jean-Christophe

APPLICANT: LIN, Winston

TITLE OF INVENTION: GENES ENCODING SEVERAL POLY (ADP-RIBOSE) GLYCOHYDROLASE (PARG)

TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACT

TITLE OF INVENTION: THEREWITH

FILE REFERENCE: NIAD 201

CURRENT APPLICATION NUMBER: US/09/511,477

CURRENT FILING DATE: 2000-02-23

PRIOR APPLICATION NUMBER: 09/302,812

PRIOR FILING DATE: 1999-04-30

NUMBER OF SEQ ID NOS: 38

SEQ ID NO 38

LENGTH: 29793

TYPE: DNA

ORGANISM: Caenorhabditis elegans

FEATURE:

US-09-511-477-38

Query Match 13.2%; Score 31.4; DB 4; Length 29793;

Best Local Similarity 51.0%; Pred. No. 7.8; Indels 0; Gaps 0;

Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 94 GATGTTCTCGAAAAGGAGCTTCCCAAGTTATTGGAGTACTGTGAAGAGTTCGTCATGAA 153

DB 21188 GATTTCTACATCTCTTATGCGCAAGCATGTTTTCATATGATGAAATTTCTGTT 21129

QY 154 GTTTACCCCAAGGACITTTACTATGTGAATTAATTTCAAACTAGTAGTCAGATCAATAA 213

DB 21128 TTTTCCCATTAAGTGAATTCATAATTTTATGTCATATGCTTGGTCTAGTAAAT 21069

QY 214 AATCTAGTCGCAAAAAA 238

DB 21068 TTTAATTTTGAATAAATAA 21044

RESULT 8

US-09-511-507-38/c

Sequence 38, Application US/09511507

Patent No. 6395543

GENERAL INFORMATION:

APPLICANT: JACOBSON, Myron K.

APPLICANT: JACOBSON, Elaine L.

APPLICANT: AME, Jean-Christophe

LENGTH: 29793

TYPE: DNA

ORGANISM: Caenorhabditis elegans

FEATURE:

US-09-302-812-38

Query Match 13.2%; Score 31.4; DB 4; Length 29793;

Best Local Similarity 51.0%; Pred. No. 7.8; Indels 0; Gaps 0;

Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 94 GATGTTCTCGAAAAGGAGCTTCCCAAGTTATTGGAGTACTGTGAAGAGTTCGTCATGAA 153

DB 21188 GATTTCTACATCTCTTATGCGCAAGCATGTTTTCATATGATGAAATTTCTGTT 21129

QY 154 GTTTACCCCAAGGACITTTACTATGTGAATTAATTTCAAACTAGTAGTCAGATCAATAA 213

DB 21128 TTTTCCCATTAAGTGAATTCATAATTTTATGTCATATGCTTGGTCTAGTAAAT 21069

QY 214 AATCTAGTCGCAAAAAA 238

DB 21068 TTTAATTTTGAATAAATAA 21044

RESULT 7

US-09-511-477-38/c

Sequence 38, Application US/09511477

Patent No. 6337282

GENERAL INFORMATION:

APPLICANT: JACOBSON, Myron K.

APPLICANT: JACOBSON, Elaine L.

APPLICANT: AME, Jean-Christophe

APPLICANT: LIN, Winston

TITLE OF INVENTION: GENES ENCODING SEVERAL POLY (ADP-RIBOSE) GLYCOHYDROLASE (PARG)

TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACT

TITLE OF INVENTION: THEREWITH

FILE REFERENCE: NIAD 201

CURRENT APPLICATION NUMBER: US/09/511,477

CURRENT FILING DATE: 2000-02-23

PRIOR APPLICATION NUMBER: 09/302,812

PRIOR FILING DATE: 1999-04-30

NUMBER OF SEQ ID NOS: 38

SEQ ID NO 38

LENGTH: 29793

TYPE: DNA

ORGANISM: Caenorhabditis elegans

FEATURE:

US-09-511-477-38

Query Match 13.2%; Score 31.4; DB 4; Length 29793;

Best Local Similarity 51.0%; Pred. No. 7.8; Indels 0; Gaps 0;

Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 94 GATGTTCTCGAAAAGGAGCTTCCCAAGTTATTGGAGTACTGTGAAGAGTTCGTCATGAA 153

DB 21188 GATTTCTACATCTCTTATGCGCAAGCATGTTTTCATATGATGAAATTTCTGTT 21129

QY 154 GTTTACCCCAAGGACITTTACTATGTGAATTAATTTCAAACTAGTAGTCAGATCAATAA 213

DB 21128 TTTTCCCATTAAGTGAATTCATAATTTTATGTCATATGCTTGGTCTAGTAAAT 21069

QY 214 AATCTAGTCGCAAAAAA 238

DB 21068 TTTAATTTTGAATAAATAA 21044

RESULT 8

US-09-511-507-38/c

Sequence 38, Application US/09511507

Patent No. 6395543

GENERAL INFORMATION:

APPLICANT: JACOBSON, Myron K.

APPLICANT: JACOBSON, Elaine L.

APPLICANT: AME, Jean-Christophe

```

APPLICANT: LIN, Winston
TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOSYLASE (PARG) EN
TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
TITLE OF INVENTION: THEREWITH
FILE REFERENCE: NIAID 201
CURRENT APPLICATION NUMBER: US/09/511,507
CURRENT FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/302,812
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 38
LENGTH: 29793
TYPE: DNA
ORGANISM: Caenorhabditis elegans
FEATURE:
US-09-511-507-38

Query Match 13.2%; Score 31.4; DB 4; Length 29793;
Best Local Similarity 51.0%; Pred. No. 7.8;
Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 94 GATGTTCTCGAAGAGCTCCCAAGTTATTGGAGTACTGTGCAAGAGAGTTCGTGTCATGAA 153
Db 21188 GATTTCTCATCTCTTATCCCAAGTCTTTCATATGTTTATGAAATTTCTTGT 21129

QY 154 GTTATCCCAAGAGCTTACTATGTGTAATTAATTTGTAACACTAGTACAGATCAATAA 213
Db 21128 TTTTCCCATTAAGTGAATTCAAATTTTATTTGTCATATGCTTGGTTCTAGTAAAT 21069

QY 214 AATCTAGTGGCAAAAAA 238
Db 21068 TTTAAATTTTCAAAAAAAGAA 21044

RESULT 9
US-09-227-357-69/c
Sequence 69, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P203CP1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795

```

```

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 69
LENGTH: 696
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (605)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (648)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (655)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-69

Query Match 13.1%; Score 31.2; DB 4; Length 696;
Best Local Similarity 50.7%; Pred. No. 2.2;
Matches 72; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 95 ATGTTCTCGAAGAGCTCCCAAGTTATTGGAGTACTGTGCAAGAGTTCGTGTCATGAG 154
Db 278 ATGCTTTTAAAAATCCAAACTAACTAATATCATGATGAGCCATGCGGTTCGG 219
QY 155 TTATCCCAAGAGCTTTACTATGTGAATTAATTTGTAACACTAGTACAGATCAATAAA 214
Db 218 GTAACTAATCCCTATACACTGTTATTTTATTATTATTATTATTATTATTATTATT 159
QY 215 ATTTACGTGGCAAAAAA 236
Db 158 TTTCACCTGCAAAAAAAGAA 137

```

```

RESULT 10
US-09-601-198-149
; Sequence 149, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Elson Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 149
; LENGTH: 2244
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-143
Query March 12.9%; Score 30.8; DB 4; Length 2244;
Best Local Similarity 55.7%; Pred. No. 4.5;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 90 TTCGATGTTCTCGAAGAGGACTCCCAAGTATTGAGTACTGTGAAGAGTTCGTCA 149
D6 338 TTAACATTACATTAAACACCAATAACAAGTTATTATTAATACATAGCTGTTTTT 397
QY 150 TGAAGTTTACCCAAAGCACTTTACTATGTGCAATTAATTTGTCAAC 195
D6 398 TTAACAGTAACAAACAACAACTATCTACAGAATTAATAATAAAAC 443

RESULT 11
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2822)..(2822)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g

```



```

? LOCATION: {657081}..(657081)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {657203}..(657203)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {67435}..(67435)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {682442}..(682442)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {713652}..(713652)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {741684}..(741684)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {779455}..(779455)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {779676}..(779676)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {855539}..(855539)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {871619}..(871619)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1084830}..(1084830)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1096846}..(1096846)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1119881}..(1119881)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1130881}..(1130881)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1310988}..(1310988)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1313224}..(1313224)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1349473}..(1349473)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1349491}..(1349491)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1470091}..(1470091)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1569020}..(1569020)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1602912}..(1602912)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1603734}..(1603734)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1637998}..(1637998)
? OTHER INFORMATION: n equals a, t, c, or g
? NAME/KEY: misc_feature
? LOCATION: {1664854}..(1664855)
? OTHER INFORMATION: n equals a, t, c, or g

```

US-09-437-450a-40.rni

```

Query Match 12.8%; Score 30.6; DB 4; Length 1664976;
Best Local Similarity 45.6%; Pred. No. 56; Mismatches 129; Indels 0; Gaps 0;
Matches 108; Conservative 0;

QY 1 TTTTCTTTTGGGAGGAAATCAGCGCGGGGATGCAACAGTCTCTCTCAATGGCA 60
DB 619908 TTATAATTACTGACATAGATACCTGTGATATTTAAATAGGACTTTCCGAGAGTAA 619849

QY 61 ACTGTCTATTATCCATTCGGCAATCAGTTTGGAGTTCCTCGAAGAGACTTCCCAAG 120
DB 619848 ATTATATTTTATTATCATTTATGATAGGACGTAATTTATATGATGATGATCAAG 619789

QY 121 TTATTCGAGTACTGCGAAGAGTTGCTCATGAGTTTACCCAAAGGACTTTTATGTGA 180
DB 619788 GGTGTAAGTTCGAATAGACTTTTGTAGTTTAAAGCATTAGTCACTTTTGTGC 619729

QY 181 ATTAAATTTGCAAACTAGTACTGAGTCAATCAATATAATTTCTAGTGGCAAAAAA 237
DB 619728 AATGATAATTTTGTCTACGAACTTCTATCAGAGATGATATTTATAAAACTACA 619672

RESULT 12
US-09-437-597-266/c
; Sequence 266, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasi: A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.45SCLL
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 266
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(401)
; OTHER INFORMATION: n = A,T,C or G
US-09-437-597-266

```

```

Query Match 12.8%; Score 30.4; DB 4; Length 401;
Best Local Similarity 49.7%; Pred. No. 3.1;
Matches 76; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 83 ATCATTTTGGAGTTCTCGAAGAGACTTCCCAAGTATTGAGTACTGGAAGAG 142
DB 181 ATTACAAAATGATTGTTTAAAGAAATTTTAAATCCGAAAGTCAATTTAAATAG 122

QY 143 TTCGTCATGAGTTTACCAAGGACTTTTACTATGTGAATTAATTTCTCAACTAGT 202
DB 121 AACCTCATATAGTATGACAACTATAAAATTTTACATTTTCAAAATTTTCTGT 62

QY 203 CAGATCAATAAAATTTCTAGTGGCAAAAAA 235
DB 61 CGTCTCATTTATTCGTCGTATACAAATTGAA 29

```

US-09-437-597-266/c

; Sequence 266, Application US/0948084A

; Patent No. 6482597

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C6

; CURRENT APPLICATION NUMBER: US/09/480,884A

; CURRENT FILING DATE: 2001-08-27

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 266

; LENGTH: 401

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(401)

; OTHER INFORMATION: n = A,T,C or G

US-09-437-450a-266

Query Match 12.8%; Score 30.4; DB 4; Length 401;

Best Local Similarity 49.7%; Pred. No. 3.1;

Matches 76; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 83 ATCACAATTCGGATGTTCTCGAAAAGAGACTTCCCAAGTATTCGGAGTACTGTGAAGAG 142

DB 181 ATTACAAAATGATTGTTAAGAAATTTTAAATCCAGAAAGTCATTAAATAG 122

QY 143 TTGTCATGAGTTTACCCAAAGAGCTTCTACTATGTAATTAATTTGCAAACTAGTAGT 202

DB 121 AACCTCATATAGTAGAACAACTATAAAATATTTTCAATTCATGAATAATTTTGTGT 62

QY 203 CAGATCAATAAAATTCACGTGCGCAAAAAAAA 235

DB 61 CGTCTCATTTATATTCNCTGTATACAAATTCGA 29

RESULT 14

US-09-542-615A-266/c

; Sequence 266, Application US/095426-5A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosker, Nancy A.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542,615A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 266

; LENGTH: 401

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(401)

; OTHER INFORMATION: n = A,T,C or G

US-09-542-615A-266

Query Match

12.8%; Score 30.4; DB 4; Length 401;

Best Local Similarity 49.7%; Pred. No. 3.1;

Matches 76; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 83 ATCACAATTCGGATGTTCTCGAAAAGAGACTTCCCAAGTATTCGGAGTACTGTGAAGAG 142

DB 181 ATTACAAAATGATTGTTAAGAAATTTTAAATCCAGAAAGTCATTAAATAG 122

QY 143 TTGTCATGAGTTTACCCAAAGAGCTTCTACTATGTAATTAATTTGCAAACTAGTAGT 202

DB 121 AACCTCATATAGTAGAACAACTATAAAATATTTTCAATTCATGAATAATTTTGTGT 62

QY 203 CAGATCAATAAAATTCACGTGCGCAAAAAAAA 235

DB 61 CGTCTCATTTATATTCNCTGTATACAAATTCGA 29

RESULT 15

US-09-606-421B-266/c

; Sequence 266, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606,421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 266

; LENGTH: 401

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(401)

; OTHER INFORMATION: n = A,T,C or G

US-09-606-421B-266

Query Match 12.8%; Score 30.4; DB 4; Length 401;

Best Local Similarity 49.7%; Pred. No. 3.1;

Matches 76; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 83 ATCACAATTCGGATGTTCTCGAAAAGAGACTTCCCAAGTATTCGGAGTACTGTGAAGAG 142

DB 181 ATTACAAAATGATTGTTAAGAAATTTTAAATCCAGAAAGTCATTAAATAG 122

QY 143 TTGTCATGAGTTTACCCAAAGAGCTTCTACTATGTAATTAATTTGCAAACTAGTAGT 202

DB 121 AACCTCATATAGTAGAACAACTATAAAATATTTTCAATTCATGAATAATTTTGTGT 62

QY 203 CAGATCAATAAAATTCACGTGCGCAAAAAAAA 235

DB 61 CGTCTCATTTATATTCNCTGTATACAAATTCGA 29

Search completed: July 7, 2004, 16:22:46

Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 15:01:33 ; Search time 344 Seconds
(without alignments)
3337.679 Million cell updates/sec

Title: US-09-437-450a-40
Perfect score: 238
Sequence: 1 tttttttttttgggggga.....tactggcaaaaaa 238

Scoring table: IDENTITY NUC

Gapop 10₀ , Gapext 1.0

Searched: 3163042 seqs, 241210380 residues

Total number of hits satisfying chosen parameters: 5326084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubna/US05_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubna/US04_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US03_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubna/US02_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US01_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubna/US00_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09D_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubna/US09E_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubna/US09F_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US09G_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US09H_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubna/US09I_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubna/US09J_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubna/US09K_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	37	15.5	263	13	US-10-085-783A-49714
2	37	15.5	263	13	US-10-085-783A-49714
3	36.2	15.2	4233	12	US-10-242-535A-49714
4	34.2	14.4	308	13	US-10-041-018-135
5	34.2	14.4	614	13	US-10-424-599-87084
6	34.2	14.4	614	13	US-10-027-632-94282
7	34.2	14.4	614	13	US-10-027-632-94282
8	34.2	14.4	614	16	US-10-027-632-94282
9	33.4	14.0	404	10	US-09-918-995-17073
10	33.4	14.0	499	9	US-09-783-590-5790
11	33.2	13.9	8781	15	US-10-311-455-1659
12	32.8	13.8	678	9	US-09-770-149-358
13	32.8	13.8	745	9	US-09-764-855-13
14	32.8	13.8	745	15	US-10-072-349-13

15	32.8	13.8	3423	16	US-10-398-221-3632
16	32.6	13.7	6117	15	US-10-243-485-43
17	32.4	13.6	698	13	US-10-342-887-765
18	32.4	13.6	698	13	US-10-172-118-765
19	32.2	13.5	15373	15	US-10-311-455-440
20	32.2	13.5	251364	15	US-10-175-523-58
21	32.2	13.5	251364	15	US-10-175-523-61
22	32.2	13.5	251364	15	US-10-175-523-79
23	32	13.4	2133	15	US-10-074-475-86
24	31.8	13.4	607	13	US-10-027-632-224261
25	31.8	13.4	607	16	US-10-027-632-224261
26	31.8	13.4	2000	9	US-09-938-842A-4562
27	31.8	13.4	2000	11	US-09-938-842A-4562
28	31.8	13.4	6565	13	US-09-938-842A-4562
29	31.8	13.4	96597	12	US-09-938-842A-4562
30	31.8	13.4	402850	10	US-09-844-653-5
31	31.6	13.3	506	17	US-10-021-323-5635
32	31.6	13.3	644	10	US-09-814-353-18833
33	31.6	13.3	778	15	US-10-106-698-589
34	31.6	13.3	1873	15	US-10-175-523-37
35	31.6	13.3	1887	17	US-10-437-963-45512
36	31.6	13.3	2034	13	US-09-973-278-872
37	31.6	13.3	2787	13	US-09-973-278-873
38	31.6	13.3	4259	8	US-08-927-939-30
39	31.6	13.3	16509	15	US-10-311-455-1294
40	31.6	13.3	113515	15	US-10-311-455-2147
41	31.4	13.2	502	10	US-09-918-995-20058
42	31.4	13.2	2280	16	US-10-369-493-42107
43	31.4	13.2	2322	17	US-10-437-963-44178
44	31.4	13.2	25083	12	US-09-937-722-121
45	31.4	13.2	23793	9	US-09-973-451-38

ALIGNMENTS

RESULT 1
US-10-085-783A-49714
; Sequence 49714, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49714
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-49714

Query Match 15.5%; Score 37; DB 13; Length 263;
Best Local Similarity 52.8%; Pred. No. 0.77;
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 56 TGGCACTGCTATTATCCATTCGCGATCACATTCGGATTCCTCGAAAGACTCC 115
DB 49 TGTAAAGATTGATTAGTACCTCATTTTCTTAATGAGTTTAAAGTTTG 108
QY 116 CAAAGTATTGAGTACTGTGAAGAGTTGTCATGAGTTTACCAAGACTTTTACTA 175
DB 109 CTTTCCCTTGATAGATTGTACATCATGCTCCCAAGATATATCCAAATTTACTA 168

```
QY 176 TGTGAATTAATTTGTCACAACTAGTAGTCAGATCAATA 212
Db 169 TAAAGATTAATTTTAAACCTTTATGTTAAATTACA 205

RESULT 2
US-10-242-535A-49714
; Sequence 49714, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49714
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-49714

Query Match 15.5%; Score 37; DB 16; Length 263;
Best Local Similarity 52.2%; Pred. No. 0.77;
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 56 TGGCACTGCTATTATTCGATTCGCGAATCACATTCGGAGTGTCTCGAAGAGGACTTCC 115
Db 49 TGGTAAGATGATGATTAGTACTTACCTCAATTTTATCTTAATGGAGTTTAAAGTAAGTTG 108

QY 116 CAAAGTATTGGAGTACTGTGAAGAGTGTGTCATCACTTTACCCAAAGGACTTTACTA 175
Db 109 CTTTGCCTTGATGAAATGTTAAACATCACTGCTCCAGAAATATTAATTCATTAATTTACTA 169

QY 176 TGTGAATTAATTTGTCACAACTAGTAGTCAGATCAATA 212
Db 169 TAAAGATTAATTTTAAACCTTTATGTTAAATTACA 205

RESULT 3
US-10-041-018-135
; Sequence 135, Application US/10041018
; Publication No. US20040072323A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Hart, Elizabeth A.
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
; FILE REFERENCE: P02080US/10025547
; CURRENT APPLICATION NUMBER: US/10/041,018
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259880
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 135
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
US-10-041-018-135

Query Match 15.2%; Score 36.2; DB 12; Length 4233;
Best Local Similarity 55.0%; Pred. No. 4.8;
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
QY 110 ACTTCCAAAAGTTATTGGAGTACTGTGAAGAGTTTCGTATGAAGTTTACCCAAAGGACT 169
Db 236 ACTTCCAGTTTTCATAAAGATAGGATGGAATTTCTAAATTAACCTTTTAACTCATA 295

QY 173 TTAATATGTAATTAATTTGTCACAACTAGTAGTCAGATCAATAAATTTCTACGTGGCAA 229
Db 296 TAAGTAATGGATTAAATTTCCGAAATTAATAAACAGAAAAATTAATTTCAAAAGTATGA 355

QY 230 AAAAAAAA 238
Db 356 AATAAAAA 364

RESULT 4
US-10-424-599-87084
; Sequence 87084, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yonqwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 87084
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: DAT_MRT3847_49647C.1
US-10-424-599-87084

Query Match 14.4%; Score 34.2; DB 13; Length 308;
Best Local Similarity 60.0%; Pred. No. 5.6;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 144 TCGTCATGAGTTTACCCAAAGGACTTTACTATGTGAATTAATTTGTCAAACCTAGTAGTC 203
Db 57 TTGTGATGTTGTTAGAAATATATTTTACCATGTGACGACTATCTCTGACATTTGTT 116

QY 204 AGATCAATAAATTTCTACGTGGCAAAAAA 238
Db 117 ATTACAAATTTTTCATGTGTCAAAAA 151

RESULT 5
US-10-027-632-94282/c
; Sequence 94282, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-03-28
```

RESULT 7
US-10-227-632-94282/c
; Sequence 94282, Application US/10027632
; Publication No. US20030204075A9

```

RESULT 8
US-10-027-632-305384/c
? Sequence 305384, Applicant:cn US/10027632
? Publication No. US20030204075A9
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
? Polymorphisms in the Human Genome
? FILE REFERENCE: 108827.129
? CURRENT APPLICATION NUMBER: US/10/027,632
? PRIOR FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218,006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 305384
? LENGTH: 614
? TYPE: DNA
? ORGANISM: Human

```

US-09-703-590-5790/c
Sequence 5790, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1

```

; NAME/KEY: misc feature
; LOCATION: (396)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (408)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (417)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (423)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (439)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (456)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (467)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (469)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (470)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (476)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (487)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (492)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (497)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-5790

Query Match 14.0%; Score 33.4; DB 9; Length 499;
Best Local Similarity 50.3%; Pred. No. 12;
Matches 76; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 66 CTATTATCCCATCCCAATCATTTCCGATGTTCTCGAAAGGACTTCCCAAGTTATT 125
   |||||
Db 28: CTACTATACNCCCAATNAATTAATTAATCTACTACGAGGTTNCCACTTTTCA 222
   |||||

QY 126 GGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTGANTAA 185
   |||||
Db 221 TAAGTTCTTANCAACATTTTCATTTTGGTTATTTCAGAACATCTCAACAGCATGTCTAA 162
   |||||

QY 186 ATTCTCAACTAGTAGTCAGATCAATAAAT 216
   |||||
Db 161 ATGTTCAATTAATAAAGGTAATTACATGAAT 131
   |||||

RESULT 11
US-10-311-455-1659/c
; Sequence 1659, Application US/10311455

; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete-
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1659
; LENGTH: 8781
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-1659

Query Match 13.9%; Score 33.2; DB 15; Length 8781;
Best Local Similarity 55.1%; Pred. No. 52;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 121 TTATTCGAGTACTGTGAAGAGTTCGTGATGAAGTTTACCCAAAGGACTTTACTATGTGA 180
   |||||
Db 191 TTATAATAAATCTTATTATCTACTATATAATAAATCTTTATATAATAAATCTTTATTATACTA 132
   |||||

QY 181 ATTAAATTCGCAAACTAGTCAGATCAATAAATTCAGTGCACAAAGGAAAAA 238
   |||||
Db 131 ATTATAATCAACAATATATTAAACATATAAACAATAATACCTTCTAATAAACTAAA 74
   |||||

RESULT 12
US-09-770-149-358/c
; Sequence 358, Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurlban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2024 (PAPA-013PRV)
; CURRENT APPLICATION NUMBER: US/09/770,149
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,506
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 678
```

; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-149-358

Query Match 13.8%; Score 32.8; DB 9; Length 678;
Best Local Similarity 64.5%; Pred. No. 22;
Matches 48; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 163 AAGACCTTACATGTCGAATTAATGTCAACTAGTAGTCAGATCAATAAATCTCAG 222
DB 83 ATGTAATGTTTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 24

QY 223 TGGCAAAAAAAAAA 238
DB 23 TAGTTAAAAAAAAAAAA 8

RESULT 13
US-09-764-855-13/c
; Sequence 13, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; CURRENT APPLICATION NUMBER: US/09/764,855
; CURRENT FILING DATE: 2002-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (56)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (397)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (512)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (587)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (711)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-855-13

Query Match 13.8%; Score 32.8; DB 9; Length 745;
Best Local Similarity 60.5%; Pred. No. 22;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 153 AGTTTACCCAAAGGACTTTACTATGCAATTAATTTCTCAACTAGTAGTCAGATCAATA 212
DB 108 AGATGATTTTAGGGACATTTTATTTTAAATTAATTTTACAATCTAATGGTNANAAAAA 49

QY 213 AAATTTCTACGTGGCAAAAAAAAAA 238
DB 48 AAAAAAAAAAAAAAAAAAAAAA 23

RESULT 14
US-10-072-349-13/c

; Sequence 13, Application US/10072349
; Publication No. US20030054420A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P110C1
; CURRENT APPLICATION NUMBER: US/10/072,349
; CURRENT FILING DATE: 2002-02-11
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 13
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (56)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (397)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (512)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (587)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (711)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-072-349-13

Query Match 13.8%; Score 32.8; DB 15; Length 745;
Best Local Similarity 60.5%; Pred. No. 22;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 153 AGTTTACCCAAAGGACTTTACTATGCAATTAATTTCTCAACTAGTAGTCAGATCAATA 212
DB 108 AGATGATTTTAGGGACATTTTATTTTAAATTAATTTTACAATCTAATGGTNANAAAAA 49

QY 213 AAATTTCTACGTGGCAAAAAAAAAA 238
DB 48 AAAAAAAAAAAAAAAAAAAAAA 23

RESULT 15
US-10-398-221-3632
; Sequence 3632, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3632
; LENGTH: 3423
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b


```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3632

Query Match      13.8%; Score 32.8; DB 16; Length 3423;
Best Local Similarity 59.8%; Pred. No. 44;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY      58  GCAACTGCTATTATCCATTCGCAATCAATTCGGATGTTCTCGAAGAGGACTTCCCA 117
Db      2377  GAAATTGTTTATTCTTTTCCCAAGCHCAAAAGCAATGGTATGGATGAAACTTCGCT 2436

QY      118  AAGTTATGGAGTACTGTGAAAGAGTTCCTCA 149
Db      2437  CAACATCGGCATAGTAGCAAAATGTTTGGCA 2468
```

Search completed: July 7, 2004, 16:28:39
Job time : 345 secs